Transportation Research Part B: Methodological, Volume 134, April 2020, pp. 229-255
DOI:10.1016/j.trb.2020.02.012

A Framework for Collaborative Air Traffic Flow Management Minimizing Costs for Airspace Users: Enabling Trajectory Options and Flexible Pre-tactical Delay Management

Yan Xu[a,b], Ramon Dalmau[b], Marc Melgosa[b], Adeline Montlaur[b], Xavier Prats[b]

[a] *Centre for Aeronautics, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, MK43 0AL, Bedford (UK)*
[b] *Department of Physics - Aeronautics Division, Technical University of Catalonia, Castelldefels 08860, Barcelona (Spain)*

Abstract

This paper proposes a collaborative air traffic flow management (ATFM) framework, in the scope of trajectory based operations, aiming to improve the cost-efficiency for airspace users (AUs) when facing ATFM regulations. The framework consists of four modules. The first one involves the AUs initially scheduling their preferred trajectories for all their flights. Based on this initial demand, the second module (assumed to be on the Network Manager -NM- side) detects time-varying hotspots (i.e. overloaded sectors along the day). In the third module, hotspot information is shared back to the AUs who plan alternative trajectory options to avoid crossing these congested airspace volumes (in the lateral and vertical domain); as well as providing to the NM different pre-tactical delay management preferences (including ground holding, linear holding, air holding and pre-tactical delay recovery); based on their internal cost breakdown structures. Incorporating all these potential combined options, the last module computes the best trajectory selections and the optimal distribution of delay assignments, such that the cost deviation from the initial status (all the user-preferred trajectories) is minimized. This model is formulated as mixed integer linear programming (MILP) and validated by a real-world case study focused on 24 hours of traffic over the French airspace. Results using the proposed framework suggest a significant system delay reduction by nearly 97% over the existing method, whilst yielding an average of less than 100 kg extra fuel consumption and 50 Euro extra route charges for the 11% flights diverted to their alternative trajectories.

Keywords: air traffic flow management, trajectory based operations, demand and capacity balancing, collaborative decision-making, trajectory options

1. Introduction

The air transportation system currently faces a significant strain from the fast-growing flight demand. This has been evidenced in recent years by severe flight delays and more commonly-seen network congestion in many regions across the world. For example, in Europe, year 2016 saw an average departure delay per flight of 11.3 minutes (and 29.1 minutes, per delayed flight, for the average arrival delay), an increase of 9% in comparison to 2015. Further, over this time period flights delayed more than 30 minutes increased 9.8%, with an average of 1.9% for operational cancellation monthly (EUROCONTROL, 2017a). A series of reports, e.g., Cook and Tanner (2015), can be used as a reference for European delay costs incurred by airlines. In the United States, in turn, 17% of the flights were delayed by more than 15 minutes in 2016, with another 1.2% canceled (US Department of Transportation, 2016). Ball et al. (2010) presented the economic impact of flight delay, where the cost of delay to airlines is estimated by modeling the relationship between the total cost and operational performance metrics. Given an average delay cost of \$62.55/min anticipated for the U.S. passenger carriers (Airlines for America, 2016), the 60 million minutes of total delay in this year led to an estimated \$3.8 billion direct aircraft operating costs.

One of the primary causes for those delays and congestion is that the number of flights (demand) often exceeds the supply of the airspace accommodation (capacity). In addition, the sustained growth in traffic also shows some seasonal or exceptional peaks (holiday seasons, major sport events, etc.). Conversely, convective weather, airspace restrictions, overloaded airports and air traffic control (ATC) industrial actions, to name a few, can temporarily reduce this supply. The effort thereby to achieve demand and capacity balancing (DCB) is typically known as Air Traffic Flow Management (ATFM).

Examples of ATFM systems include the Enhanced Tactical Flow Management System (ETFMS), implemented by Eurocontrol's Network Manager Operations Centre (NMOC, previously Central Flow Management Unit), which compares traffic demand, regulated demand and load against capacity to assess possible imbalances in the European airspace and allows the implementation of measures to resolve these imbalances in the traffic, such as regulations or rerouteing. With the assistance of the Network Operations Plan (NOP) Portal, European ATFM stakeholders will have access to the up-to-date information of the network situation which will allow them to more

dynamically plan and manage the demand and capacity (EUROCONTROL, 2017b). Similar initiatives exist in the United States, including Ground Delay Programs (GDPs) and Airspace Flow Programs (AFPs). GDPs control the arrival rate at an affected airport by assigning departure delays to flights at their origin airports (FAA, 2009). An AFP identifies constraints in the en-route system, regulating flights filed into the Flow Constrained Area (FCA) (Libby et al., 2005). While a flight has no choice but to eventually end up at its destination airport, a capacity-constrained en-route sector can often be bypassed at limited cost by selecting an alternative route. To that aim, AFPs specify available reroutes that avoid the FCA. Flight operators may then choose to accept the delay for an affected flight, or to take the available reroute (Pourtaklo and Ball, 2009).

The overall objective of these ATFM initiatives is typically to reach a compromise solution across all stakeholders based on some fairness criteria (e.g., first scheduled, first served policy). In this context, specific preferences for the airspace users (AUs) were not typically taken into account in the early development of ATFM programs. With the paradigm shift for the future air traffic management (ATM) proposed by SESAR (Single European Sky ATM Research) in Europe and NextGen (Next Generation Air Transportation System) in the United States for instance, the AUs are expected to increasingly participate in ATM decisions, using, in particular, more collaborative decision making (CDM) mechanisms.

CDM is more of a philosophy for better managing air traffic through information exchange, procedural improvements, tool development, and common situational awareness (Ball et al., 2000). It allows decisions to be taken by those best positioned to make them based on the most comprehensive, up-to-date accurate information and ensuring that all concerned stakeholders are given the opportunity to influence the decision (EUROCONTROL, 2017b). CDM was first implemented in GDPs in the late 1990s (Chang et al., 2001), and then incorporated tools such as flight substitution, cancellations, compression, and slot credit substitution (Ball et al., 2005). SESAR has been advancing this through development of the User Driven Prioritisation Process (UDPP) to achieve additional flexibility for AUs to adapt their operations in a more cost-efficient manner (SESAR, 2015). Increased CDM is also found in the Collaborative Trajectory Options Program (CTOP) that has been deployed since early 2014 in the United States, and which is built upon concepts of GDPs and AFPs. AUs are allowed to submit, well in advance of the issuance of the program, a set of desired reroute options that

can be used to route around an FCA or constraint (FAA, 2014). In general, under CDM, ATFM is expected to be conducted in a way that gives significant decision-making responsibilities to AUs (Vossen et al., 2012).

To achieve the top-level aspiration of the ATM paradigm shift described above, one must overcome the amalgamation of the flight planning and execution phases, based on advanced flight trajectory management. Indeed, the flight trajectory is established as the fundamental element of such operating procedures, which is referred to as Trajectory Based Operations (TBO). Concretely, TBO represents an ATM method for strategically planning, managing and optimizing flights throughout the operation by using time-based management, information exchange, and the aircraft's ability to fly precise paths in time and space (SESAR, 2020; FAA, 2018). It requires innovations to be introduced in all parts of the ATM system to realize the envisioned changes. Stakeholder involvement, better data sharing and usage with System Wide Information Management (SWIM), introduction of advanced decision support tools for human operators, both on ground and in the air, and improving management in all the facets of the air transportation, are just a few envisioned and needed changes.

The TBO concept has nowadays became the main focus of validation conducted in International Civil Aviation Organization (ICAO) Global Air Navigation Plan (ICAO, 2016), as well as in multiple regional programmes, including the SESAR (Europe) and NextGen (U.S.), and also CARATS (Japan), OneSKY (Australia) and Sirius (Brazil). Although there may exist slight differences for the TBO-associated definitions through the different programmes, the general principle is indeed of high degree of consistency. For the sake of clarity, this paper takes all terminology from the SESAR TBO concept (SESAR, 2017).

In the above context of TBO, a framework for Collaborative Air Traffic Flow Management (C-ATFM) is introduced in this paper. The main structure of the framework is presented in Fig. 1, which is composed of four modules, each representing the tasks that might be conducted by either the AUs or the NM. An outline of each module is given as follows:

- Module I: Initial planning of user-preferred trajectories

  This module refers to the planning of trajectories by the AUs, taking into account forecast weather conditions and strategic ATM constraints, such as route availability restrictions or flight level allocation and orientation schemes (see Section 3.1). The trajectory optimization

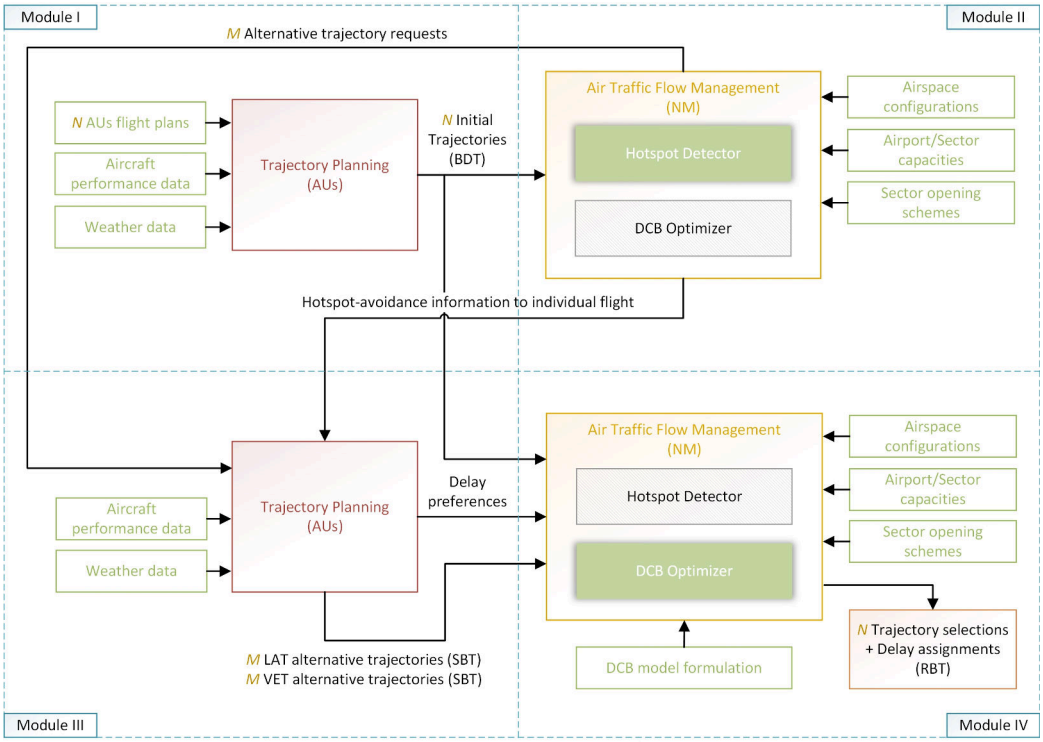


Figure 1: An overview of proposed collaborative air traffic flow management framework.

and planning methodology used in this module has been previously reported in (Dalmau et al., 2018). According to the SESAR concept of operations (ConOps), these trajectories would correspond to the Business Development Trajectories (BDT).

- Module II: Detection of demand and capacity imbalance

  Based on the trajectories computed in the previous module (i.e. initial traffic demand), a primary detection of imbalances between traffic demand and airspace capacity is conducted in this module (see Section 3.2). Time-varying hotspot volumes are thereby identified. Combined with airspace geometric descriptions, the specific hotspot avoidance information is shared back to all AUs with one or more concerned flights, i.e., flights traversing at least one hotspot (see Section 3.3).

- Module III: Submission of trajectory options and pre-tactical delay management preferences

With the hotspot avoidance information received, concerned AUs compute alternative trajectories for their captured flights to avoid entering these hotspot volumes, using the same trajectory optimization techniques implemented for initial trajectory planning (see Section 3.4). Different preferences are also allowed on how AUs wish to manage (at pre-tactical/dispatch level) the delay (see Section 3.5). According to the SESAR ConOps, these trajectories would correspond to the Shared Business Trajectories (SBT).

- Module IV: System-wide optimization to balance demand and capacity

  This module is initiated by the NM to balance the demand and capacity, yielding eventually the best combinations of trajectory selections and delay assignments among all regulated flights (see Section 4). The objective considered in this paper minimizes the overall deviation with respect to the ideal status where all AUs could maintain their initial BDTs. Resulting trajectories would correspond to the Reference Business Trajectories (RBT).

The modular design of this framework allows flexible adjustment of one or more modules for various purposes. In particular, this paper focuses on the benefits analysis with respect to an ideal deterministic setting for the framework over the current operations. Taking into account the uncertainty in the system, stochastic models could be adopted in Module IV to better address such issue in reality. Also, additional mechanisms (such as UDPP) for collaborative trajectory planning can be considered in Module III to give more priority to the equity concern. Then, a real-world case study is presented in Section 5, using realistic data of 24 hours traffic crossing the French airspace. Finally, Section 6 summarizes the conclusions of this paper. The supplement materials are provided in Appendix from A through E.

2. Literature review

Following the pioneering work done by Odoni (1987), researchers have focused on the development of models to minimize the congestion costs in response to airport capacity reduction. Delay assignment, such as ground holding, has been used as the most common short-term ATFM initiative (see (Terrab and Paulose, 1992; Richetta and Odoni, 1994) for instance). If

congestion at airspace sectors is also taken into account, the problem of controlling release times and speed adjustments as well as reroutings of aircraft while airborne for a network of airports (including sectors) was studied in (Bertsimas and Patterson, 1998, 2000; Lulli and Odoni, 2007). With the added complication of the problem, dynamical rerouting proved highly effective in the case of weather affected approaches around the airport which itself can operate at full capacity (Mukherjee and Hansen, 2009). Aiming at realistic applications, the computational challenges arising from previous models were largely reduced by means of a massively parallel Dantzig-Wolfe decomposition method applied to the Bertsimas Stock-Patterson model (Rios and Ross, 2010; Tandale et al., 2013).

Regarding CDM, researchers have explored different ways to incorporate mechanisms to previous models, aiming at further improving ATFM performance. Moreover, some potential metrics to achieve an acceptable fairness level, compensated by some loss of system efficiency, were proposed and discussed in (Barnhart et al., 2012; Bertsimas and Gupta, 2015). An efficient dual network flow formulation for the static-stochastic GDP was presented in (Ball et al., 2003), showing how this formulation can be implemented under CDM with equity considerations. The integer programming formulation for the GDP was extended by Vossen and Ball (2006a), to approximate the CDM process where the slot compression step can be considered as a mediated bartering between AUs. The opportunities for slot trading in a single airport setting were studied under the condition that GDP offers are given to trade from various airlines (Vossen and Ball, 2006b). Similarly, slot exchange mechanisms in an AFP scenario through a mediated bargaining of assigned slots was discussed in (Sherali et al., 2011), allowing AUs to improve cost-efficiency. The overall collaboration process was then simulated by Molina et al. (2014) using an agent-based modeling approach, where different ATM stakeholders were modeled in a CDM framework.

For the current version of CTOP, an RBS (ration-by-schedule) scheme is adopted. Namely, flights are assigned the best available routes and slots available at the time flight operators submit their preference requests during the planning period, in a sequential manner (Miller and Hall, 2015). Yet, the rules of allocation in that algorithm have some obvious drawbacks, such as airlines' competitive responses. In this context, Kim and Hansen (2015) investigated a game theoretic treatment of airline preference submission behavior within the First Submitted First Assigned allocation process. Recently, an alternative flight scheduling approach for CTOP, based on lin-

ear optimization instead of pure RBS, has been studied using a Max-Min fairness rule to guarantee some equity (Rodionova et al., 2017).

Furthermore, to enhance the involvement in the CDM paradigm, there has been also much research conducted from the AUs' point of view. Early studies (Meyer and Oster, 1981; Morrison and Winston, 2010) explored the impacts of deregulation and the statistical relationships between airline operating cost variables and financial performance, with a focus on fuel and crew costs. Holloway (2008) provided an overview of the different types of schemes established to categorize costs in the airline operations. Flight cancellation decisions in GDP were studied in (Xiong and Hansen, 2009), revealing the value of a flight cancellation and airline preference structure in their decision making process.

With the forthcoming TBO concept, a transition in ATM from control by tactical clearance to management by reference to a trajectory is expected, which emphasizes the importance of an efficient trajectory planning (optimization) from AU side. The aircraft trajectory optimization problem can be modeled and solved using optimal control theory (Betts and Cramer, 1995; Betts, 2010). A recent study (Gardi et al., 2016) provided a comprehensive review of the different trajectory optimization techniques, with a special focus on the recent advances introduced in the ATM context. Using optimal control has the advantage that the dynamic equations of the aircraft are taken into account, producing very accurate and realistic trajectories while obtaining the guidance commands as result of the optimization process. Complete frameworks using this approach have been reported in (Soler et al., 2012; Dalmau et al., 2018). Using this technique, Xu et al. (2017); Xu and Prats (2017a) showed how to optimally handle ATFM and additional reactionary delays at dispatch (planning) level by means of linear holding and pre-tactical delay recovery strategies.

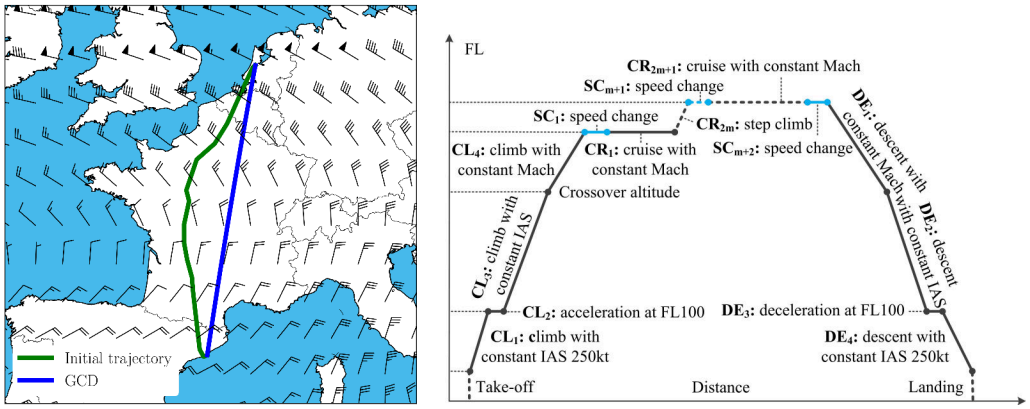
3. Collaborative Trajectory Design

This section introduces the interactive trajectory design process, aligned with the CDM paradigm described before. Specific avoidance information is generated by the NM and is shared to concerned AUs for each affected flight. Precisely-designed alternative trajectories, along with preferences on delay management, are produced by AUs and eventually submitted back to the NM (Modules I, II and II in Fig. 1).

3.1. Initial schedule of user-preferred trajectory (Module-I)

In the European ATFM system, AUs have been offered a high level of flexibility in regard to flight planning (Bolić et al., 2017), which enables the scheduling of initial trajectory to well reflect their preferences. For day-to-day operations, these preferences are generally focused on the aircraft direct operating cost encompassing a weighted sum of fuel consumption, route charges and time-related costs (such as crew and maintenance fees).

Aiming at minimizing the total operating cost per flight, in this paper both the lateral route and vertical profile are optimized to generate an optimal 4D trajectory in order to represent the user-preferred trajectory (i.e. the BDT). Nevertheless, the details of producing this initial traffic demand are out of the scope of the paper and the methodology was previously reported in (Dalmau and Prats, 2017; Dalmau et al., 2018). Next, a brief summary of this methodology is given.



(a) Lateral route affected by weather (b) Vertical profile with ATM restrictions

Figure 2: Initial trajectory planning decoupled to lateral route and vertical profile.

The trajectory optimization algorithm used here decouples the computation of the optimal lateral route and vertical profile. The available lateral route network is represented by a graph, in which the nodes represent waypoints or navigation fixes and the edges are the route segments. Based on this graph, the optimal lateral route minimizing the direct operating cost is computed by using the A* algorithm (Hart et al., 1968). The difference in national unit rates with regards to the route charges has been taken into account. Realistic weather conditions, especially wind fields that have a great impact on the trajectory, are also considered, using GRIdded Binary

(GRIB) formatted files (World Meteorological Organization, 1994). The effects of wind on the optimal lateral route can be seen in Fig. 2a, in which GCD (Great Circle Distance) represents the shortest distance, whereas the actual optimal lateral route follows the green line's path.

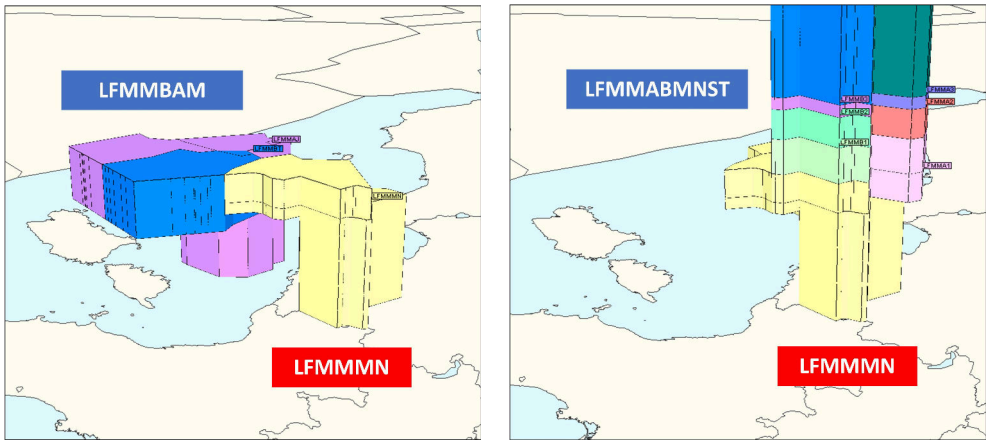
The vertical profile is modeled by a given sequence of parametrized flight phases. Each phase within the flight profile contains information about the aerodynamic configuration and throttle setting, and may also include constraints representing AUs' operations and ATM restrictions, as shown in Fig. 2b. Then an optimal control problem is solved to obtain the best vertical profile, minimising a given cost function and subject to these constraints. To obtain accurate fuel consumption and time figures, aircraft performance data from the Base Of Aircraft Data (BADA) v4 (Nuic and Mouillet, 2014) published by Eurocontrol are used, along with the above mentioned weather data (e.g., temperature and wind field).

In addition, it is assumed in this paper that AUs can provide their BDTs at the beginning of the day and that these remain valid throughout the day. In some instances, delays earlier in the day can propagate so that later BDTs may need to be time shifted. Under current operations, these newly filed flights will be treated as a popup, and will be assigned with the delay received by other flights planned to enter the affected area at about the same time. However, this simple (yet fair) rule may not apply directly to the case in this study, as the impacts of delaying two similar flights (even with the same O/D pair and crossed sectors) could be quite different in a network scenario. To tackle this issue, an online appendix introduces some possible ways of multi-stage decision making, which will allow AUs to update the BDTs (and others) before the progressive decisions have been made on their flights.

3.2. Detection of time-varying hotspot airspaces (Module-II)

On basis of the initially planned trajectories, a primary detection on the imbalances of traffic demand and airspace capacity is conducted by the NM, identifying the hotspot volumes. Under the trajectory based operations, it is clear that not only airspace capacities vary with time, due to sectorization schemes or weather changes for instance, but also traffic demand, which depends on the scheduled flights for that day and the particular realization of 4D trajectories for each flight. Thus, the detected hotspot airspaces are also time-varying.

The capacity of an airport can be evaluated (and quantified) in different ways such as number of take-offs/landings per unit of time, runway occupancy time or terminal entry/exit rate. Similarly, for airspace sectors the traffic demand could be counted by means of aircraft entry rate, occupancy, density or complexity. For simplicity, the entry rate is adopted as the criterion to count sector demand, which is also the method commonly used in current operations. Capacity values for this paper are directly retrieved from the Demand and Data Repository v2 (DDR2) database published by Eurocontrol (EUROCONTROL, 2018). It is worth noting that the approach proposed in this paper would also be applicable to other definitions of demand and capacity via certain adjustments (if needed).



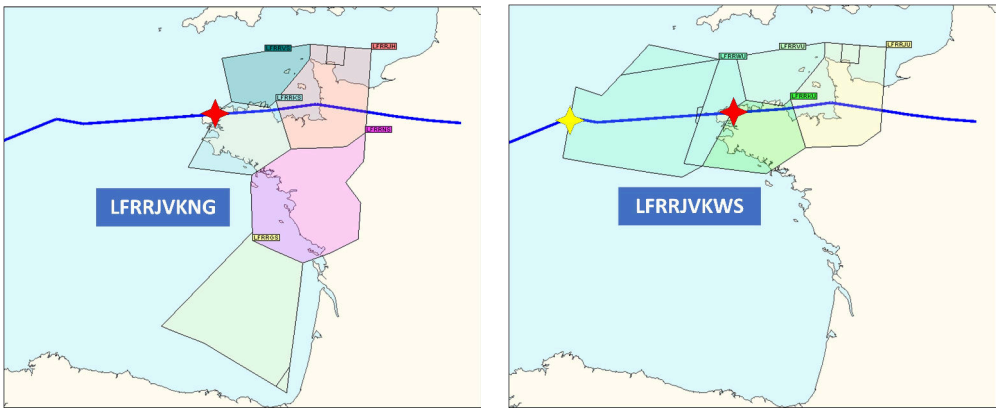
(a) Collapsed sector LFMMBAM (b) Collapsed sector LFMMABMNSR

Figure 3: Elementary sector LFMMMN collapses to different operating sectors during different time periods of the same day.

Specifically, as for counting the entry rates, the (flight) entry points (into any airspace volume) are defined at the boundary of elementary sectors, as their geographical dimensions normally remain stable during a relatively long period, e.g., an AIRAC (Aeronautical Information Regulation And Control) circle of 28 days. However, an elementary sector could be collapsed with its adjacent elementary sector(s) in a much shorter time scale, acting as operating sector as a whole. For example, as illustrated by Fig. 3, an elementary sector LFMMMN (colored in yellow) is merged into two different collapsed sectors (LFMMBAM and LFMMABMNSR) in two different hours of the same day, along with other elementary sectors. Obviously, the two

newly-formed operating sectors have quite different physical dimensions, and in many cases, they would have different operating capacities.

The difference between elementary sector (to define entry point) and collapsed sector with realistic operating capacity, requires a specific judgment about whether a flight entry shall be counted as a traffic demand. If an aircraft enters a collapsed sector such as LFRRJVKNG in Fig. 4a, having an intersection with a particular elementary sector that belongs to this collapsed sector, then, according to the above discussion, the intersection is always seen as a flight entry (and is also subject to a control point as will be discussed in Sec. 4). However, as shown in Fig. 4b, for exactly the same entry point (labeled with a red star), assume that the intersected elementary sector belongs to another collapsed sector (e.g., LFRRJVKWS) at another time when the flight actually enters (because the flight has been delayed). In this case, the particular flight entry should not be counted as an extra traffic demand of that operating sector, which is, instead, treated as an internal movement inside the sector.



(a) Aircraft entry (red) counted (b) Aircraft entry (red) not counted

Figure 4: Criteria of whether the same aircraft entry to elementary sector is counted as a traffic demand for different (collapsed) operating sectors (opened in different time periods).

Following this, only the first entry to the collapsed sector, e.g., the red label in Fig. 4a or the yellow label in Fig. 4b, is the one that will be counted as a demand (within a certain time period) among all the flight entries defined at various belonged elementary sectors. The above principle is given in Algorithm 1, where ETO and CTO are short for, respectively, the Estimated and Controlled Time Over the entry point.

Algorithm 1 Count traffic demand for collapsed sectors

```
1: for c in collapsed_sector_list do                          ▷ collapsed sector c
2:     for t in time_period_list do                           ▷ time period t
3:         if c in operating_sector_list[t] then              ▷ see if c is open in t
4:             for e in elementary_sector_list[c] do          ▷ ele. sector e belonging to c
5:                 for f in flight_list[e] do                 ▷ flight f traversing e
6:                     Ent = min(ETO_list[f][e])              ▷ first time of f entering c
7:                     if Ent in t then                       ▷ see if ent is in t
8:                         ini_demand[c][t]+=1                ▷ add 1 to initial demand
9:                         Cot = CTO[f][e]                    ▷ controlled time of f entering e
10:                        if Cot in t then                   ▷ see if cot is in t
11:                            reg_demand[c][t]+=1            ▷ add 1 to regulated demand
```

As can be seen from the 4th and 6th lines of Algorithm 1, there is an implicit relationship between the elementary sector (e), collapsed sector (c) and time variable (t). Namely, an elementary sector will belong to a specific collapsed sector at a certain time. As presented with Algorithm 2 in Appendix B, an additional procedure is conducted to establish such (static) scheme[1] from the published DDR2 database.

Finally, based on the counted initial traffic demand and the retrieved airspace capacity, hotspot sectors are detected as a function of the time. Typically, the time scale for capacity evaluation is 20 min or 60 min, so the time-varying hotspot positions may evolve among the entire airspace network after every 20 min or 60 min. This list of predicted hotspots is the basis of the avoidance information provided to each individual concerned flight, as discussed in the next section.

3.3. Hotspot avoidance information to individual flights (Module-II)

With the time-varying hotspots detected, all flights that are planned to traverse those airspace volumes during the corresponding time period will be captured. Concerned AUs will be inquired to submit alternative trajectories

[1]One implicit problem is that the given airspace structures, such as the schemes of collapsing/splitting sectors, are often designed to best accommodate the planned (or historical) demand. Once trajectories have been changed, by imposing delay or rerouting, the temporal-spatial traffic flow patterns (and thus hotspots) will accordingly change, and consequently the initial airspace structures may be not optimal. To solve this issue, a follow-up study (Xu et al., 2018) presented a method to realize optimizing traffic flow and scheduling airspace configuration in a synchronized way.

(to avoid hotspots) for each of their affected flights. This submission is not mandatory for the AUs, who could decide that for some (or all) of their flights the initial trajectory is the only option, which will be likely subject to (significant) delay. Hence, AUs will have to consider the extra costs incurred if flying alternative trajectories (possibly subject to less or no delay) and decide whether the submission of alternatives is worthwhile based on the cost-breakdown particularities of each concerned flight.

It is not necessary to require a single flight to evade all of the identified hotspots. There could be multiple areas identified during the same period of time across the entire network of airport and sectors. In fact, only the hotspot(s) that a particular flight traverses according to its initial trajectory (i.e. the BDT) have to be bypassed by providing (if desired by the AU) alternative trajectories. This is because hotspots change with time, as discussed in Sec.3.2, and entry times of the (eventual) alternative trajectories (i.e. the SBTs) can only be known by the NM once the AU have submitted them. Once all this information is gathered, then the NM will compute a solution that respects all capacity constraints in all airspace sectors and airports (see Section 4).

Given that airspace sectors include lateral coordinates of boundary points and vertical altitudes of lower and upper bounds, a flight, in theory, should be able to avoid a sector in both lateral and vertical domains (except for those sectors close to the departure or arrival airport). For convenience, they are entitled hereafter as lateral-avoidance and vertical-avoidance alternative trajectories respectively. Note that the cruise altitude(s) may be different than the initial when applying the lateral avoidance, which is determined by the optimal vertical profile based on the diverted lateral route.

To assist AUs to design these lateral and vertical avoidance trajectories for their concerned flights, some specific information is shared with them. An example taken from Eurocontrol DDR2 is given in Table 1, where it is shown a hypothetical flight captured by a hotspot in sector LFEEKDF. This sector is a collapsed sector as a consequence of the merging of two elementary sectors (LFEEKF and LFEELD), which in turn are constructed by a set of airblocks. In DDR2, airblocks are the elemental airspace structures containing geographical information (i.e. coordinates for the vertices).

For lateral-avoidance, the vertex coordinates of each airblock are given in such a way that a specific polygon graph can be formed on the horizontal plane to represent the entire hotspot area. For vertical-avoidance, based on the initial trajectory, it informs for each sector that the flight needs to avoid,

Table 1: Precise avoidance information shared for an individual flight to allow the AU to plan lateral and vertical alternative trajectories avoiding concerned hotspots.

	Airblock* code	Total vertices	Vertex 1		Vertex 2		Vertex n	
			Lat. (min)	Long. (min)	Lat. (min)	Long. (min)	Lat. (min)	Long. (min)
Lateral avoidance	200LF	27	2920	187	2922	210	...	...
	202LF	35	2920	187	2922	210	...	...
	203LF	36	2920	187	2922	210	...	...
	204LF	16	2895	344	2879	360	...	...
	Sector code	Entry dist. (nm)	Exit dist. (nm)	Lower alt. (100 ft)	Upper alt. (100 ft)			
Vertical avoidance	LFEEKF	-196	-117	345	999			
	LFEELD	-234	-196	345	375			

* In DDR2, elementary sectors are defined by the union of several airblocks.

at which distance to the destination airport (e.g., -234 nm) the flight should start to change the original altitude and at which other distance (e.g., -196 nm) the vertical avoidance is not longer needed. The information shared also specifies the non-selectable flight levels (e.g., from FL345 to FL375) between these two distances.

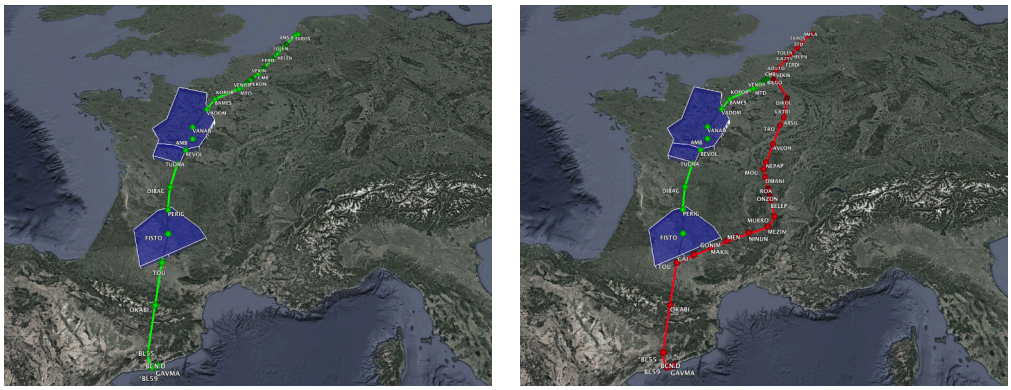
This hotspot-avoidance information could be taken into account by AUs to generate trajectory options, and the next section will introduce how to make use of it effectively from their perspective. In addition, it is worth noting that the reason of providing such set of accurate data is for reducing, as much as possible, the extra costs yielded from diverting a flight to its alternative trajectory, which accounts for a key performance of the system-wide optimization model as discussed later in Sec. 4.

3.4. Lateral and vertical avoidance alternative trajectories (Module-III)

After receiving the detailed hotspot avoidance information, AUs could generate the alternative trajectories for their affected flights, with the same tools used for planning the initial trajectories (refer to Sec. 3.1), whilst adding additional constraints (as specified within the avoidance information).

As shown in Fig. 5a, for the same flight introduced in Fig. 2a (Barcelona El Prat - Amsterdam Schiphol), the initial trajectory (green line) has been captured, since it is scheduled to traverse two hotspot sectors. Once the AU receives the avoidance information for these two hotspots, the lateral avoidance trajectory is computed by means of removing from the graph all those edges crossing any of the boundaries of the sectors (recall Table 1). Then, a re-computation of the (optimal) vertical profile is triggered by the AU on basis of the new lateral route (see red line in Fig. 5b). Note that

if either the origin or destination airport is inside or close to any hotspot volume, the lateral avoidance trajectory may not be possible.



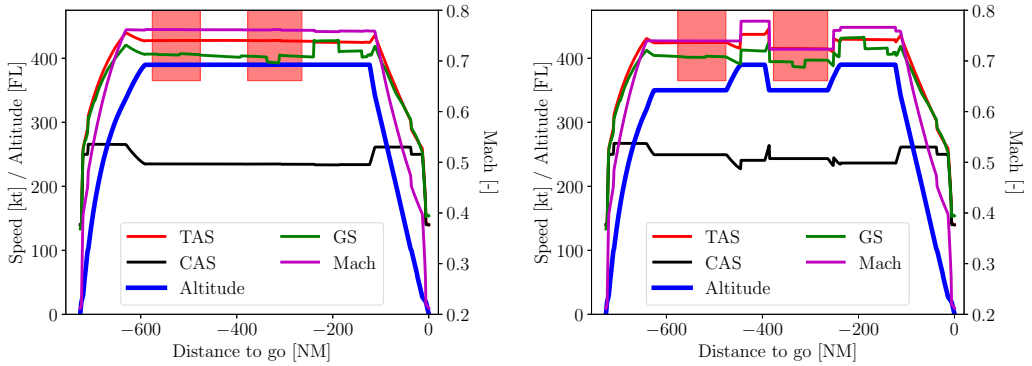
(a) Initial optimal route

(b) Alternative optimal trajectory avoiding hotspots in the lateral domain

Figure 5: 4D trajectory optimization: optimal route planning over a conventional structured ATS network

For the vertical-avoidance alternative trajectory, the lateral route is fixed to that initially scheduled and for each hotspot sector i, the segments of route traversing the hotspot are identified, together with lower and upper altitudes defining that sector, h_L^i and h_U^i, as shown by the red squares in Fig. 6a. Then, during the numerical integration of the climb phase, if the along-path distance is included into any of the climb segments, it is checked whether the aircraft would penetrate altitude h_L^i from below at the next integration step. If so, a level off at constant altitude and CAS (or Mach if h_L^i is above the cross-over altitude) would be performed until reaching the end of the segment. Then, the climb is resumed until reaching the Top of Climb (TOC) at the optimal cruise altitude (compare the TOC positions in Fig. 6a and in Fig. 6b). The same principle applies for the integration of the descent phase. When generating the cruise phase, from the TOC to the TOD (Top of Descent), the flight levels in the range $[h_L^i, h_U^i]$ are removed from the candidate set of flight levels within the segments.

At each integration step, the optimal cruise altitude (in terms of the direct operating cost) is computed to decide whether a step climb (see Fig. 6b) should be performed or not. The solving algorithm follows an iterative process similar to those implemented in state-of-the-art on-board Flight



(a) Initial optimal altitude and speed profiles

(b) Alternative optimal trajectory avoiding hotspots in the vertical domain

Figure 6: 4D trajectory optimization: optimal altitude and speed profiles under conventional flight level allocation and orientation constraints

Management System (FMS), which systematically evaluates all potential sequences of decision parameters and selects the optimal one. It should be noted, however, that eventual ATC restrictions might restrict the number of climbs/descents in cruise to avoid hotspots (to avoid the so called "yoyo" effect). Although these restrictions could be easily incorporated in the proposed methodology, for the sake of generality this paper assumes that AUs will always submit their optimal trajectories, regardless of how many flight level changes resulted from the hotspot avoidance.

3.5. Pre-tactical delay management preferences (Module-III)

In addition to alternative trajectory options, delays might still be required to solve the demand and capacity imbalance. Different types of initiatives may apply to absorb (or recover) the assigned delays in the pre-tactical phase (i.e. at flight dispatch level), and their costs, limitations and implementations are not necessarily the same. In this paper, four specific ways to handle delays at dispatch level are considered including ground holding, airborne holding, linear holding and delay recovery (Xu and Prats, 2017b). These initiatives will change the Controlled Times Over (CTOs) at positions, which can be regarded as different ways to adjust the 4D trajectory's timeline.

Airborne holding would consume more fuel due to the extended flight track, whilst ground holding has no impact on fuel consumption. Due to

the increased extra fuel, the airborne holding time is fairly limited, taking into account that safety related issues may arise from a reduction of the on-board reserve fuel. Ground holding, however, can only be performed at the departure airport, prior to take-off. Airborne holding (including holding patterns or path stretching) can be done at any available airspace, in theory, but practically it is typically performed in designated locations.

Linear holding and pre-tactical delay recovery are performed airborne too, but rather than extending the flight path, they are executed following the original trajectory by means of a cost-based speed control method (i.e. decreasing or increasing speed to absorb or recover delay, respectively). Generally, the amount of linear holding and delay recovery that can be achieved depends on factors such as the aircraft type, trip distance, payload, cruise flight level and etc., as well as the extra fuel allowance (if any) when applying these strategies. For more information, see (Xu et al., 2017; Xu and Prats, 2017a).

In order to avoid confusion, it is worth emphasizing that in this paper delay recovery is considered at flight dispatch level (pre-tactical recovery), while obeying all the ATFM controlled times of arrival (CTAs) distributed along the trajectory. Obviously, recovery can only be performed if CTOs are given in one or several positions along the route, leaving some margin to recover delay with the flight segment going from the position of the last CTO to the destination airport. Consequently, delay recovery cannot be performed if a CTO is given at the destination airport. The delay recovery considered in this paper must be differentiated from the tactical delay recovery procedures typically performed in current operations, which is a consequence of enforcing only the Controlled Time of Departure (CTD) instead of CTO/CTA at the affected sector/airport (where the latter is actually effective).

Although recovering delay could benefit the system in general, this pre-tactical delay recovery is allowed only if some delay will be imposed at the forepart of a trajectory (e.g., ground holding at the origin airport) during the delay assignment stage. The reason is as follows: this paper assumes that the BDTs/SBTs (submitted by the AUs) are their most preferred trajectories, which means that when they compute the speed profiles, the time-related costs (including the buffer time for ground operations) should have been considered already. In other words, increasing (or reducing) aircraft speed in order to recover (or absorb) delay airborne may incur some extra operating costs for particularly the AUs. Due to the lack of information in their preferences of delay recovery over the extra costs, this paper assumes, for

simplicity, that they are willing to pay more to fly faster only if their flights will be delayed. Otherwise, they would prefer to fly as initially scheduled.

4. Demand and Capacity Balancing

In this section, a linear optimization model is presented. It incorporates all potential options, such as alternative trajectories and pre-tactical delay management preferences coming from the collaborative trajectory design process presented in Section 3, in such a way that traffic demand is balanced with available capacity (Module IV in Fig. 1). The mathematical formulation is based on the well-studied Bertsimas Stock-Patterson model as presented in (Bertsimas and Patterson, 1998), which has shown excellent computational performance to handle this type of problems.

In a more recent study, Bertsimas et al. (2011b) introduced some new constraints that force local routing conditions sufficient to perform the rerouting function efficiently. However, compared with the method proposed in this paper, the rerouting decision was made in a more centralized way, which means that the concerned AUs would have to divert their flights to any possible routes specified by the NM. As discussed previously, such requests may not be favored by the AUs. One reason could be that the diverted trajectory is of low efficiency in terms of operating costs, as the NM typically lacks some proprietary information from AUs that is critical in trajectory planning. Recent studies (see (Liu et al., 2018) for instance) have been also devoted to capturing AUs' trajectory choices based on historical data, so that the ATFM solutions will be able to propose alternative routes that are most likely to be accepted by specific AUs. This paper, in light of the principle of the TBO, allows (and encourages) AUs to resubmit alternative trajectories on their own decisions, through a collaborative process as introduced in Sec. 3. In other words, the AUs always compute the trajectories, starting from the BDT (nominal flight plan), down to the RBT. The NM does not compute trajectories, but chooses the "best" option in a system-wide optimization. This distributed decision-making framework is regarded as the main advance with respect to the previous work done in (Bertsimas et al., 2011b).

4.1. Problem statement

In this paper we assume conventional airspace management, where the different ANSPs might change the configuration of the airspace according to

a finite list of different possible sectorisations (i.e. different ways to define collapsed sectors from the elementary sector list). These sectorisation changes might be done at given time intervals along the day (typically at every 20 or 60 minutes), producing the so called sector opening scheme. Typically, ANSPs might choose the best sector opening scheme that better matches the forecast traffic demand along the day, trying to avoid as much as possible sector overloads. Yet, some other operational considerations are taken into account, such as ATC staff availability, for instance. It is out of the scope of this paper to consider advanced solutions, such as the dynamic airspace configuration (DAC) concept currently explored by SESAR (Zelinski and Lai, 2011); or integrated solutions such as those proposed in (Xu et al., 2018).

The different trajectory options and pre-tactical delay management preferences resulting from the collaboration process described in the previous section are integrated into a single optimization model, selecting eventually the best distribution of trajectory options and delay assignments. This means that for each flight there are N_f feasible combinations to solve the demand and capacity imbalance, with:

$$N_f = N_t(2^{N_d} - 1) \tag{1}$$

where N_t represents the number of trajectory options submitted for flight f, and N_d is the number of pre-tactical delay management preferences that the AU envisages for the same flight. $N_t \geq 1$ since the nominal trajectory initially scheduled is always an option and $N_d \geq 1$ since ground holding must always be applicable. Then, depending on the AU, more or less trajectory or delay preferences might be submitted. It is worth noting that the -1 in Eq. (1) is needed to discard the combination where only delay recovery is performed (and no ATFM delay is assigned).

The model is formulated as mixed integer linear programming (MILP), and the corresponding decision variables are defined below. Note that detailed notations of the formulation can be found in Appendix A.

- Decision variables for trajectory options:

$$w_k^f = \begin{cases} 1, & \text{if trajectory } k \text{ is chosen for flight } f \\ 0, & \text{otherwise} \end{cases}$$

- Decision variables for delay management:

$$x_{k,t}^{j} = \begin{cases} 1, & \text{if trajectory } k \text{ departs from position } j \text{ by time } t \\ 0, & \text{otherwise} \end{cases}$$

$$y_{k,t}^{j} = \begin{cases} 1, & \text{if trajectory } k \text{ arrives at position } j \text{ by time } t \\ 0, & \text{otherwise} \end{cases}$$

Fig. 7 presents the trajectory timeline versus designed positions (i.e., intersections with elementary sectors, along with origin and destination airports), where the four delay preferences are implemented. Note that an alternative trajectory implies a new set of intermediate designed positions (e.g., P-1, P-2 and P-3 in Fig. 7).

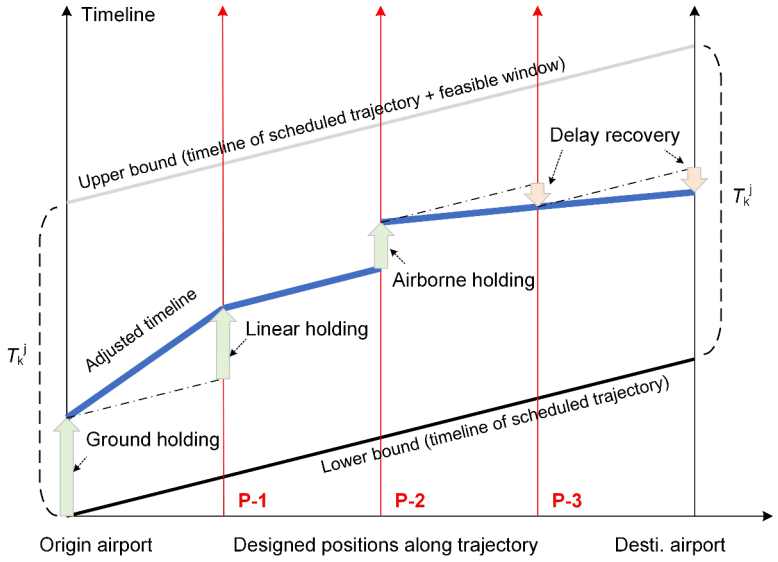


Figure 7: Schema of trajectory timeline versus designed positions.

Concretely, ground holding is experienced only at the origin airport; airborne holding can only be performed "at" a given position (the difference between the "departure" and "arrival" time at that position equals to the holding time); and since linear holding and delay recovery are realized by speed control, the slope of the lines is increased or decreased compared with the initially planned schedule.

Recall that the typical airborne holding is distinguished from linear holding in Sec. 3.5 by the fact that when performing the former, the actual flight distance will be extended (either by vectoring or using holding patterns).

This flight path "stretching", however, does not contribute to the execution of a trajectory defined by contiguous points. Thus, the typical airborne holding, on some level, can be seen as a "circling" at a particular position. It is also worth noting that the positions referred here may not correspond to the actual geographical (navigation) waypoints existing in current airspace. In addition, the feasible time window $\mathcal{T}_k^j$ shown in Fig. 7 defines a solution, based on the initially scheduled times of a particular trajectory, which will largely reduce the number of variables needed for the optimization.

Finally, note that the "by" time is used, rather than "at" for the time domain, which would enable a faster solution searching process according to (Bertsimas and Patterson, 1998), while the "at" time can be derived by $(x_{k,t}^j - x_{k,t-1}^j)$ and $(y_{k,t}^j - y_{k,t-1}^j)$ respectively. To enforce that only one time slot will be assigned to one trajectory at each designed position, within a prescribed feasible window $\mathcal{T}_k^j$, it has to satisfy $\sum_{t\in\mathcal{T}_k^j}(x_{k,t}^j - x_{k,t-1}^j) = 1$ and $\sum_{t\in\mathcal{T}_k^j}(y_{k,t}^j - y_{k,t-1}^j) = 1$. However, when using the "by" time, this constraint can be rewritten and fixed prior to solving the model, by means of setting $x_{k,\overline{\mathcal{T}}_k^j} = 1$, $y_{k,\overline{\mathcal{T}}_k^j} = 1$, and $x_{k,\underline{\mathcal{T}}_k^j-1} = 0$, $y_{k,\underline{\mathcal{T}}_k^j-1} = 0$, where $\underline{\mathcal{T}}_k^j$ and $\overline{\mathcal{T}}_k^j$ are respectively the lower and upper bound of the solution search window $\mathcal{T}_k^j$.

4.2. Objective function

In light with the discussions in Sec. 3.1, the initially scheduled trajectory should represent the most preferred trajectory for a specific flight from the AUs' point of view (i.e. individual optimum). In this way, if every single initial trajectory were maintained as it is in the final execution of flights, a global optimum at system level would be achieved, which is equal to the combination of all individual flight optima. Nevertheless, some regulations on those trajectories might be enforced due to certain reasons (e.g., a demand and capacity problem), meaning that not all the individual optima can be attained.

The objective function used in the model presented in this paper, therefore, aims to minimize such deviations[2], namely the extra fuel consumption,

[2]It is worth noting that for the reason of respecting the fairness principle, the objective function could be reformulated to minimize the maximum deviation of each flight, namely the Max-Min rule, in which no single flight can increase its benefits without reducing the benefits of other flights (Bertsimas et al., 2011a).

extra route charges, and the extra time related costs, between the initial trajectory and that trajectory resulted from the DCB process:

$$\min(C_{\Delta F} + C_{\Delta R} + C_{\Delta T}) \tag{2}$$

The summed three items correspond to those considered for the scheduling of initial trajectories (recall Sec. 3.1), which in turn are treated as the baseline for computing the extra costs. Accordingly, the extra fuel consumption $C_{\Delta F}$ is denoted as in Eq. 3.

$$C_{\Delta F} = \sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} \alpha_k (F_k - F_0^f) w_k^f \tag{3}$$

where F_k and F_0^f are respectively the total fuel consumed with trajectory k and with the initial trajectory of flight f. Note that $k \in \mathcal{K}_f$, where k is one of the trajectory options $\mathcal{K}_f$ submitted by flight f. In this case, if the initial trajectory is eventually selected, then the extra cost incurred from the fuel consumption for flight f would be equal to zero. Similarly, the extra route charges $C_{\Delta R}$ are computed by Eq. 4, where R_k and R_0^f are the total ATS fees charged with trajectory k and with the initial trajectory of flight f. α_k and β_k are respectively the weighting cost of fuel and route charges, which in turn could be specified on a trajectory basis.

$$C_{\Delta R} = \sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} \beta_k (R_k - R_0^f) w_k^f \tag{4}$$

It should be noted that, due to practical reasons, the route charges are nowadays paid based on the GCD between the entry and exit positions within the (different) charging zones based on the planned (initially scheduled) trajectory. They are not re-charged for the added and/or reduced distances caused from tactically altering the trajectories (Delgado, 2015). However, this may not reflect well the real air traffic services that they use. Therefore, aiming at future TBO concept of precise operations, this paper calculates the route charges, for each flight, using the absolute distances along the flown trajectory inside the charging zones according to the corresponding national unit rates[3]. Moreover, this is the most generic formulation of the model

[3]From 1 January 2020, Eurocontrol is now using the actual route flown as recorded by the NM to establish the distance factor used for the calculation of route charges.

presented in this paper and $C_{\Delta R}$ could always be set to zero if applying the current charging policy.

Eq. (5) defines the time-related costs discussed in Sec. 3.5, which are composed of those incurred from the different types of delay, including ground holding GH_k, air holding AH_k (i.e., standard airborne holding and linear holding), and delay recovery DR_k. It can be noticed from Eq. (5) that the cost weights (i.e., γ_k^g, γ_k^a and γ_k^d) of each delay preferences could be trajectory-specific[4].

$$C_{\Delta T} = \sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} [\gamma_k^g GH_k + \gamma_k^a AH_k - \gamma_k^d DR_k] \tag{5}$$

Since delay recovery $DR_k = GH_k + AH_k - AD_k$, where AD_k denotes the arrival delay at the destination airport, Eq. (5) can be organized as follows:

$$C_{\Delta T} = \sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} [(\gamma_k^g - \gamma_k^d) GH_k + (\gamma_k^a - \gamma_k^d) AH_k + \gamma_k^d AD_k] \tag{6}$$

Specifically, depending on the holding positions, GH_k, AH_k, and AD_k are formulated by the respective decision variables as depicted by Eqs. (7)-(9):

$$GH_k = \sum_{t \in \mathcal{T}_k^j, j = \mathcal{J}_k(i): i = 1} (t - r_k^j)(x_{k,t}^j - x_{k,t-1}^j) \tag{7}$$

$$AH_k = \sum_{t \in \mathcal{T}_k^j} \sum_{j \in \mathcal{J}_k(i): 1 < i < n_k} t(x_{k,t}^j - x_{k,t-1}^j - y_{k,t}^j + y_{k,t-1}^j) \tag{8}$$

$$AD_k = \sum_{t \in \mathcal{T}_k^j, j = \mathcal{J}_k(i): i = n_k} (t - r_k^j)^{1+\epsilon} (y_{k,t}^j - y_{k,t-1}^j) \tag{9}$$

Taking into account the fairness factor of delay assignment, the total delay is multiplied by a coefficient $(t - r_f^k)^{1+\epsilon}$ (with ϵ slightly greater than 0) in Eq. (9), in such a way that the delays would be assigned moderately across all the flights (Bertsimas and Patterson, 1998), instead of unevenly to one particular flight.

[4]For example, one particular trajectory might be given higher priority through setting a greater γ_k^d enabling more delays to be recovered. However, this may raise issues of eventual unfair competition, such as gaming. How to prevent these issues is still under research and it is out of the scope of this paper.

4.3. Constraints

The constraints concerned with the model are categorized into four groups, including aircraft operations, user-specified limits, network capacities, and decision variables.

4.3.1. Aircraft operations

$$\sum_{k \in \mathcal{K}_f} w_k^f = 1 \quad \forall f \in \mathcal{F} \tag{10}$$

$$x^j_{k,\underline{\mathcal{T}}^j_k - 1} = y^j_{k,\underline{\mathcal{T}}^j_k - 1} = 0 \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall j \in \mathcal{J}_k \tag{11}$$

$$x^j_{k,\overline{\mathcal{T}}^j_k} = y^j_{k,\overline{\mathcal{T}}^j_k} = w_k^f \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall j \in \mathcal{J}_k \tag{12}$$

$$x^j_{k,t} - x^j_{k,t-1},\ y^j_{k,t} - y^j_{k,t-1} \geq 0 \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall j \in \mathcal{J}_k, \forall t \in \mathcal{T}_k^j \tag{13}$$

$$x^j_{k,t} - y^j_{k,t} \leq 0 \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall j \in \mathcal{J}_k, \forall t \in \mathcal{T}_k^j \tag{14}$$

Constraint (10) enforces that only one trajectory of all submitted options (including the initial and alternatives) is eventually selected for each flight. Constraints (11)-(12) guarantee that each selected trajectory k, (i.e., under the condition of $w_k^f = 1$, otherwise if $w_k^f = 0$ then all decision variables associated with the unselected trajectory are equal to 0), is assigned with only one time slot for departing and arriving respectively at position j within the prescribed time window $\mathcal{T}_k^j$. Constraints (13) ensures the timeline's continuity, namely if an aircraft arrived/departed at time $t-1$ then it must have arrived/departed at time t. Constraint (14) specifies that the departure time is not earlier than the arrival time for any aircraft at any position.

4.3.2. User-specified limits

$$\begin{aligned} &y^{j'}_{k,t+u_k^{j,j'} z_k^{j,j'}} - x^j_{k,t} \leq 0 \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall i \in [1, n_k - 1] : \\ &j = \mathcal{J}_k(i), j' = \mathcal{J}_k(i+1), \forall t \in \mathcal{T}_k^j \cap [\underline{\mathcal{T}}^{j'}_k - u_k^{j,j'} z_k^{j,j'}, \overline{\mathcal{T}}^{j'}_k - u_k^{j,j'} z_k^{j,j'}] \end{aligned} \tag{15}$$

$$x^{j}_{k,t+v_k^{j,j'}z_k^{j,j'}} - y^{j}_{k,t} \geq 0 \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall i \in [1, n_k - 1]: \\ j = \mathcal{J}_k(i), j' = \mathcal{J}_k(i+1), \forall t \in \mathcal{T}_k^j \cap [\underline{\mathcal{T}}_k^j, \overline{\mathcal{T}}_k^j - v_k^{j,j'} z_k^{j,j'}] \tag{16}$$

Constraints (15) and (16) present the user-specified limits, which stipulate the time bounds of delay recovery and linear holding respectively, by means of incorporating a coefficient $u_k^{j,j'}$ (for flying faster, i.e., delay recovery) and another coefficient $v_k^{j,j'}$ (for being slower, i.e., delay absorption). These coefficients are based on $z_k^{j,j'}$, i.e., the segment flight time of the initially scheduled trajectory k crossing segment (j, j') that connects two conjunct elementary sectors j and j'. This information is to be shared by AUs who are willing to perform delay recovery and linear holding in the air for their specific flights (see (Xu and Prats, 2017b) for details on how to compute the time bounds, per flight, under certain operating costs), and they are set to 1 as default if such information is not provided.

4.3.3. Network capacities

$$\sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} \sum_{t \in \mathcal{T}_k^j \cap \mathcal{T}(\tau), j = \mathcal{J}_k(1)} (x^j_{k,t} - x^j_{k,t-1}) \leq C_j^D(\tau) \quad \forall j \in \mathcal{J}_a, \forall \tau \in \mathbb{T} \tag{17}$$

$$\sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} \sum_{t \in \mathcal{T}_k^j \cap \mathcal{T}(\tau), j = \mathcal{J}_k(n_k)} (y^j_{k,t} - y^j_{k,t-1}) \leq C_j^A(\tau) \quad \forall j \in \mathcal{J}_a, \forall \tau \in \mathbb{T} \tag{18}$$

$$\sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}_f} \sum_{t \in \mathcal{T}_k^j \cap \mathcal{T}(\tau), j = S(k,l,\tau)} (x^j_{k,t} - x^j_{k,t-1}) \leq C_l(\tau) \quad \forall l \in \mathcal{L}_\tau, \forall \tau \in \mathbb{T} \tag{19}$$

Constraints (17), (18) and (19) ensure that the traffic demand does not exceed the capacity of departure airport, arrival airport and airspace sector, respectively. The situation for an airport is relatively clear, while in the case of an airspace sector, the opening scheme $\mathcal{J}(l, \tau)$ has to be taken into account (recall Algorithm 2). Accordingly, operating sectors l are used for matching the demand and capacity, instead of elementary sectors j at which the control points are defined in this paper. Meanwhile, as shown with Algorithm 1, only the first aircraft entry into an operating sector among all the entered elementary sectors (if any) that belong to the operating sector during the certain time period τ, i.e., $j = S(k, l, \tau)$ in Constraint (19), is counted as a valid demand.

4.3.4. Decision variables

$$w_k^f \in \{0,1\} \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f \tag{20}$$

$$x_{k,t}^j,\ y_{k,t}^j \in \{0,1\} \quad \forall f \in \mathcal{F}, \forall k \in \mathcal{K}_f, \forall j \in \mathcal{J}_k, \forall t \in \mathcal{T}_k^j \tag{21}$$

Constraints (20) and (21) declare the binary (0-1) decision variables and the associated domains of the problem.

4.4. Model variants for uncertainty concerns

The above proposed framework is aimed at an ideal static scenario, but its potential applicability may take place in the context of uncertain demand and capacity. On the demand side, it is commonly accepted that actual stakeholders' behavior is not always fully "rational", and that demand predictions given by normative models and used by ATFM decision support systems contain significant variations with the actual one. This uncertainty is typically not quantified or considered in any meaningful way, which results in the system not being robust enough against behavioral biases. To tackle this issue, recent advances in data-driven technologies such as machine learning may help in modeling the real behavior (rational or irrational) and thus improving the demand predictability (Liu et al., 2018). The uncertainty in the demand aspect, however, is beyond the scope of this paper.

The capacity, on the other side, is also subject to various uncertain factors such as convective weather, as has been considered for strategic ATFM in (Yang, 2017). The European ATM system capacities, including airports and airspace, depend critically on the nature and extent of weather problems, which are uncertain and unfold over time. Thus, the weather impact on airport and airspace capacities tend to be categorized into scenarios. Based on that, the model presented in Module IV can be further extended. The detailed formulations for five model variants concerning such capacity uncertainty (in both stochastic and deterministic ways) are included in an online appendix.

5. Illustrative examples

This section presents the numerical experiments conducted under the proposed framework, with a real-world case focused on the French airspace. 24 hours of traffic have been gathered from historical demand on a typical

day in February 2017. All source data used in the experiments are based on the DDR2 database. The ATS structured route network is considered in the case study when planning trajectories.

5.1. Experimental setup

The original sample data involve 6,593 planned flights in total, but in some cases their initial trajectories only form a small part of intersection with the airspace sectors. Due to operational limits, this temporal intersection is typically not counted as an independent flight entry. In this study, 60 sec is regarded as the minimal time spent in a sector. After removing intersections less than 60 sec, there are 6,255 flights left. On the other hand, the total number of elementary sectors are 164 for that day, which are merged into 224 different collapsed sectors through the 24h period.

In addition, some key assumptions have been made in this case study:

i the unit time slot in the experiments is set to 1 min, while the time scale for matching demand and capacity is 20 min;

ii the maximal delay assigned to each flight is limited to 45 min;

iii the default cost for ground holding is considered as 81 Euro/min, according to the European network average cost of ATFM delay in 2014 (Cook and Tanner, 2015), while it is assumed with 90 Euro/min for air holding including the standard airborne holding and linear holding (which should incur no extra fuel consumption than initially planned);

iv the upper bound for performing linear holding is 20% of the segment flight time for all flights, referred to statistical average value reported in (Xu et al., 2017), and for delay recovery this bound is set to 10%, both of which are rounded to the greatest integer that is less than or equal to;

v the cost of delay recovery is -5 Euro/min for all flights, meaning that all the flights would be in favor of increasing certain speed (burning some extra fuel) to recover part of their previously experienced delays; and

vi the price of fuel, in line with the reference for the base scenario in (Cook and Tanner, 2015), is set to 0.8 Euro/kg.

5.2. Benchmark indicators

Table 2 presents the benchmark results, namely implementing C-ATFM in GH mode, which means that all the possible options mentioned in Sec. 3 are disabled, except for ground holding. This is similar to the CASA

(Computer Assisted Slot Allocation) function adopted currently within Eurocontrol's Enhanced Tactical Flow Management System. CASA follows the principle of RBS and matches traffic demand and airspace capacity by delaying flights' departure times (Cook, 2007). The key difference, however, is that the RBS "constraint" is not imposed for C-ATFM (GH mode). Note that in this benchmark test the maximal amount of delay per flight is limited to 480 min for the C-ATFM (GH mode), instead of 45 min specified above for the full-functional mode.

Table 2: Benchmarks of total delays and flights for CASA and C-ATFM (GH mode).

Cases	**Delay (min)**	**Flights (a/c)**	**Avg. delay (min)**	**Avg. cost (Euro)**	**Dep. rev. (#)**	**Arr. rev. (#)**
CASA	406,042	2,510	162	13,122	0	0
GH mode	220,044	1,881	117	9,477	15,424	20,606

As Table 2 shows, only about half of the delays are required by C-ATFM (GH mode) with respect to those needed by CASA. This reveals, on some level, the trade-off between efficiency (i.e., minimizing the total delay cost in GH mode) and equity (i.e., obeying the "first-come, first-served" principle in CASA). In other words, C-ATFM (GH mode) gives the minimum delay (since it optimizes the delay assignment), while CASA always assigns slots according to that equity rule. Note that the trade-off effects could be further enlarged for a greater network, as there will be even larger amount of nodes (airports and/or sectors) where the RBS rule should apply.

A certain amount of capacity overloads are usually allowed in reality (and in some cases the allowance can be quite large). This could be due to several reasons, such as the lack of initial schedules for pop-up flights, the conservative method for capacity evaluation, and the current way of counting traffic demand (i.e., flight entry rate) without considering the factors of occupancy, traffic pattern and complexity. Nevertheless, for the illustrative purpose, no capacity allowance is allowed in this study, which also accounts for the huge delays (see Table 2) assigned in these benchmark experiments that should not occur in real-world operations.

The above results suggest that RBS is inefficient, but another view is that RBS is so important that the AU community is willing to incur extra delay in order to maintain it. Such importance, in the sense of equity, could be comprehended from the perspective of flight reversals, namely the overtaking for each pair of flights. This sounds natural considering the notion of fairness

widely agreed on by the AUs is to have a schedule that preserves the order of flight arrivals/departures at an airport according to the published schedules (Bertsimas and Gupta, 2015). In this study, to achieve the half amount of delay reduction with C-ATFM (GH mode), it appears 15,424 and 20,606 pairs of overtaking for departure and arrival respectively (see Table 2).

5.3. Hotspot detection and trajectory options

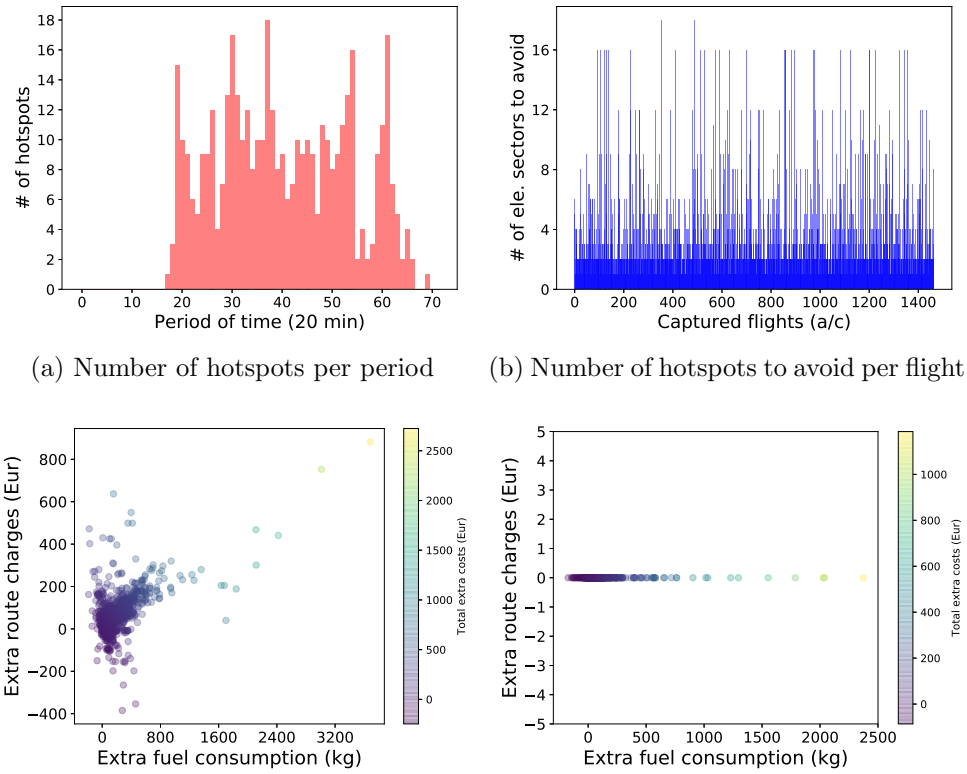
The results using C-ATFM, with full-functional mode, are presented hereafter. Through the hotspot detection process (recall Sec. 3.2), there are 433 different time-varying hotspots identified, as shown in Fig. 8a. Then, the captured 1,464 flights (that are initially scheduled to fly across any of the hotspots) are required to provide alternative trajectories making use of the sector avoidance information shared to each of them (see Sec. 3.3). The number of elementary sectors that each captured flight needs to bypass is as shown in Fig. 8b.

Subsequently, 1,305 lateral and 1,379 vertical alternative trajectories are generated by AUs and returned to the NM. The missing ones are due to the fact that some hotspot volumes may be located close to origin/destination airports and/or restricted areas and consequently can not be avoided either with lateral and/or vertical trajectory alternatives. Wrapping up, there are 8,939 (6,255 initial + 1,305 lateral + 1,379 vertical) trajectories scheduled for 6,255 flights.

The extra costs required for diverting flights to their lateral and vertical (hotspot-avoidance) alternative trajectories are given in Figs. 8c and 8d respectively. These extra costs include fuel consumption and route charges. It can be seen from Fig. 8c that, in many cases, a longer route (to avoid hotspots) between a given O/D pair may lead to lower route charges, due to the different national unit rates, yet such reduction in route charges may detract from the increased fuel consumption. The wind effects can be also appreciated in this figure, as there appear some cases where the total costs can decline to even lower than initially scheduled. Besides, it can be noticed that flying at sub-optimal altitudes may incur notable extra costs, although no additional route charges should apply (see Fig. 8d).

5.4. Case studies

From its beginning, CDM has recognized the need for a structure that encourages AUs to provide accurate information. For example, the use of RBS (for equity) and specifically the policy that AUs can retain arrival slots



(a) Number of hotspots per period

(b) Number of hotspots to avoid per flight

(c) Extra costs for lateral alternatives

(d) Extra costs for vertical alternatives

Figure 8: Time-varying hotspots and extra costs for lateral and vertical hotspot-avoidance alternative trajectories.

for cancelled flights is intended to motivate them to share flight cancellation information. As presented, the proposed framework appears to offer AUs strong incentives to misrepresent costs and search for other ways to manipulate their inputs to the system, e.g., by proposing terrible alternative trajectories the AU may increase the chances that the optimum solution will assign the initial trajectory.

To address this drawback, a set of case studies have been conducted to explore possible ways of enhancing AUs' incentive and equity in the framework. The main principle is to pose less penalty to certain flights having more potential extra costs (such as to undertake more delays or to select alternative trajectories). Each following ordered case is considered based on the previous one, incorporating new features step by step:

- Case-0: the original model using all default parameter values given in experimental setup (see Sec. 5.1);
- Case-1: Case-0 + setting superlinear factor ϵ to 0.2, to further take into account the non-linearity of delay cost for each flight;
- Case-2: Case-1 + restricting the maximal delay quota of 10 min for each alternative trajectory, but 45 min (as default) for the initial trajectory;
- Case-3: Case-2 + allowing only ground holding to be used as pre-tactical delay management initiative for the initial trajectory; and
- Case-4: Case-3 + imposing a weighted cost $1 + \frac{(\Delta F + \Delta R)}{(F+R)}$ on the delay assigned to alternative trajectory based on its extra and initial cost.

For each case, the problem dimensions and computational times are summarized in Table 3. The mathematical formulation is based on the Bertsimas Stock-Patterson Model. In numerical experiments, GAMS v.25.1 software has been used as the modeling tool and Gurobi v.8.1 MIP optimizer as the solver. Computations have been run on a 64 bit Intel i7-8700 @ 3.20 GHz six-core CPU computer with 32 GB of RAM and Linux OS.

Table 3: Problem size and computational time.

Summary	**Case-0**	**Case-1**	**Case-2**	**Case-3**	**Case-4**
Variables	4,943,940	4,943,940	3,665,250	3,665,250	3,665,250
Equations	11,343,818	11,343,818	8,334,973	7,064,986	7,064,986
Non-zeros	27,925,736	27,925,736	20,577,424	15,107,278	15,107,278
Generation	2 min	2 min	2 min	1 min	1 min
Solution	138 min	99 min	62 min	14 min	9 min
Objective	412,848	412,848	423,893	610,545	611,441
Rel. gap	0.28%	0.50%	0.42%	0.00%	0.00%

Experimental results can be found in Table 4, which include key indicators encompassing delay assignment, trajectory selection, extra costs and equity. Across all these cases, the most promising result would be that the total (arrival) delay is reduced to only 3,000 - 4,500 min. Remember that when using C-ATFM with GH mode, this number is greater than 200,000 min (see Table 2), which means that the delay reduction (by using the C-ATFM full version) is nearly 98%. In the meantime, the total delay cost is reduced by 95% (due to the cost variance of using different delay initiatives), and the

average delay cost (per delayed flight) also decreases from almost 10,000 Euro (see Table 2) to less than 600 Euro.

Table 4: Main indicators of the results derived from each case of study.

Indicators	**Case-0**	**Case-1**	**Case-2**	**Case-3**	**Case-4**
Total delay (min)	3,083	3,106	**3,137**	4,413	4,411
Delayed flight (a/c)	621	626	**633**	937	935
Avg. delay cost (Euro)	567	565	**560**	402	402
Original trajectory (a/c)	5,455	5,439	**5,460**	5,406	5,404
Alternative trajectory (a/c)	800	816	**795**	849	851
Avg. delay for original (min)	4.8	4.9	**5.2**	4.9	4.9
Avg. delay for altern. (min)	5.8	5.9	**2.8**	3.5	3.4
Extra fuel consumption (kg)	73,777	73,779	**74,886**	79,454	78,781
Extra route charges (Euro)	23,936	23,908	**23,873**	25,359	24,985
Extra trajectory cost (Euro)	60,825	60,798	**61,316**	65,086	64,376
Avg. extra cost per altern. (Euro)	76	75	**77**	77	76
Departure reversal pairs (#)	235	225	**231**	248	251
Arrival reversal pairs (#)	190	178	**174**	257	255

Next, such notable delay (and its cost) reduction is accompanied by approximately 800 flights diverted to their alternative trajectories. The extra trajectory costs (including fuel consumption and route charges) range between 60,000 - 65,000 Euro (in which the fuel cost is slightly higher than the charges of additional route), making the average extra cost per selected alternative trajectory close to 75 Euro. In other words, by means of diverting 800 flights (i.e., 13% of the total) incurring 75 Euro extra trajectory cost for each (see Table 4), the system delay will decrease by more than 200,000 minutes, producing a net income (i.e., delay cost reduction minus trajectory cost increase) of 16 million Euro in a system wide view[5]. This trade-off seems impressive, which will be further demonstrated in Sec. 5.7.

In terms of fairness indicators, the numbers of flight reversal pairs are lowered down to 200 - 300 for both departure and arrival across all the cases (which is due to the absolute delay reductions), if compared with more than 15,000 in the GH mode (recall Table 2). Further, when allowing the airborne linear holding and delay recovery, part of the reversed flights in departure can be recovered till the arrival through speed changes en route.

[5]Note that the reason of issuing such a large amount of delay (in C-ATFM GH mode) is because all capacity constraints through the network have been strictly enforced (for illustrative purpose), which is not the real case in current operations.

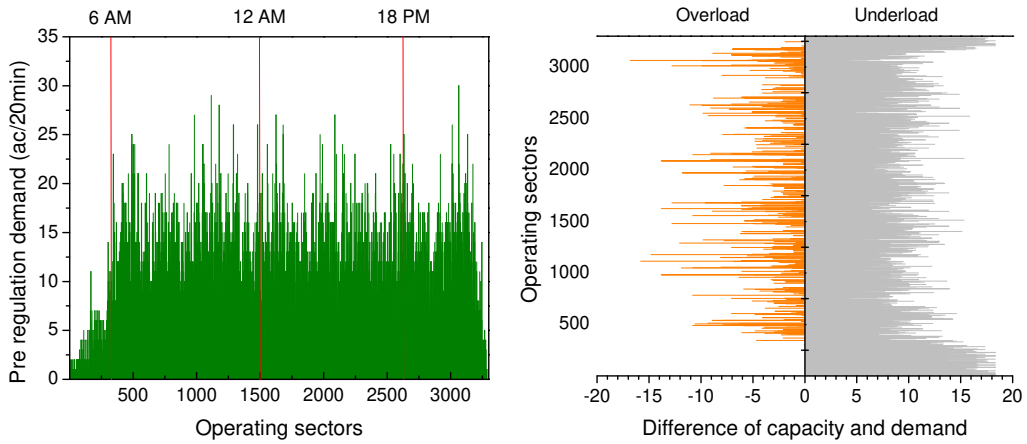
Then, for different cases of study, Case-1 poses some non-linearity in the cost of delay, which yields more flights to be delayed and diverted with an increased (average) extra trajectory cost. The consequence is, however, a decrease in flight reversals (i.e., improved equity) as shown in Table 4. From Case-1 to Case-2, each alternative trajectory is given a quota of delay, namely 10 min maximal, and hence the average delay for the alternatives is reduced from 5.9 min to 2.8 min, compared with an increased number (from 4.9 min to 5.2 min) for the originals. This feature could be interpreted as one strong incentive for AUs to submit accurate alternative trajectories, as they can be guaranteed with no more than a slight delay.

From Case-2 to Case-3 and Case-4, even less penalty is given to the flights having more potential extra costs, but in the meantime it largely increases delay (i.e., 40% more than that for Case-2). This means that some additional efficiency is lost for achieving a higher level of equity (and incentive). In this sense, Case-2 might be relatively a balanced method, and therefore is chosen as the way for conducting all the following experiments in this study.

Summing up, this section discusses the concerns for AUs' incentive and equity. These two factors are indeed the key aspects of mechanism design, which deserve a thorough assessment in future work. The main focus of this paper is limited to quantify the benefit pool of the framework that might be attained with an ideal system operating (in a deterministic setting). This ideal system would include an incentive compatible mechanism for eliciting the required inputs from AUs, and not be constrained by equity concerns.

5.5. Overall demand and capacity situations

Fig. 9a presents firstly the initial (i.e., pre-regulation) demand for each considered operating sector during each time period. The total number of operating sectors across the day (72 periods of time) is 3,285, each of which is formed in a period of 20 min by either an elementary sector (164 in total) or a collapsed sector (224 in total). The sequence of these operating sectors has been ordered in accordance with their activation time in that day (but is arbitrary within each 20 min period). It can be also noticed that there are more sectors opened from 6 AM to 6 PM than the reverse, which is also roughly in line with the distribution of traffic occurrence. Seeing from Fig. 9b, large numbers of capacity overloads (i.e., demand higher than capacity) can be found, while in some cases it could be as high as twice the capacity value that the sector can provide.



(a) Initial demand in operating sector (b) Imbalance of demand and capacity

Figure 9: Initial traffic demand and imbalances with capacity for each operating sector during each time period.

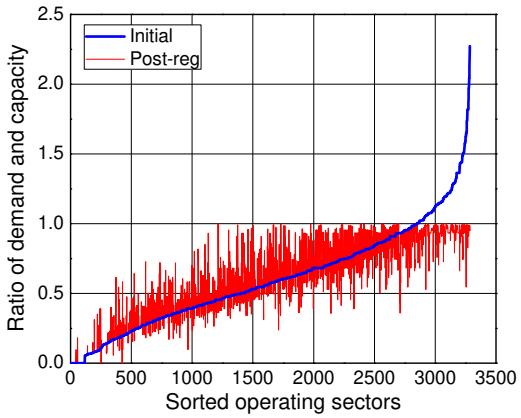


Figure 10: Sorted ratios of demand and capacity for initial and post-regulation situations.

To better understand the balance between demand and capacity, their ratios are sorted (based on the initial demand) and presented in Fig. 10. Obviously, the curves representing pre-regulation (i.e., initial) are steeper with some parts growing higher than 1 (i.e., demand higher than capacity). Conversely, the curves turn to be level and average with respect to the post-regulation cases, which means that more airspace capacities are well utilized. Note that Fig. 10 contains only data for the day of operation in 24 hours,

without showing a number of regulated flights having their arrival times delayed to the next day. In other words, the regulated demand (represented by the blue line) would be slightly lower than the initial demand (red line), where the gap is 695 (i.e., 27,601 - 26,906) entry counts (2.5% of the total) in this particular case study.

5.6. Trajectory selections and delay assignments

As mentioned previously, any of the options proposed in Sec. 3.4 and Sec. 3.5 can be integrated together and imposed on one flight. For this particular case of study $N_t = 3$ (nominal trajectory plus lateral and vertical alternatives) and $N_d = 4$ (ground holding, airborne holding, linear holding and pre-tactical delay recovery), leading to 45 possible combinations (recall Eq. 1). For example, an aircraft might be asked to experience some ground holding at the origin airport, fly its lateral alternative trajectory, undertake a small amount of airborne or linear holding en route, whilst being allowed to partially recover those delays along the remaining trajectory.

Table 5: Summary of trajectory selections and delay assignments* in C-ATFM.

Options	**Initial** (5,460 a/c)		**Lateral** (530 a/c)		**Vertical** (265 a/c)		**Total** (6,255 a/c)	
	Flight (a/c)	Time (min)	Flight (a/c)	Time (min)	Flight (a/c)	Time (min)	Flight (a/c)	Time (min)
GH	819	3,890	63	174	36	103	918	4,167
AH	24	161	2	3	1	3	27	167
LH	42	87	3	5	1	3	46	95
DR	619	-1,150	54	-100	25	-42	698	-1,292
AD	579	2,988	29	82	25	67	633	3,137

* GH-ground holding; AH-airborne holding; LH-linear holding; DR-delay recovery; AD-arrival delay

In the previous case study, there are 795 flights selecting their alternatives, which contain 530 lateral hotspot-avoidance trajectories and another 265 vertical ones, as shown in Table 5. According to the cost results presented in Figs. 8c and 8d, the vertical alternatives generally incur less extra cost than the lateral, due to lower extra fuel consumption and no additional route charges. However, it turns out that more laterals than the verticals are eventually selected. This suggests that the lateral alternative trajectories should perform better in terms of redistributing the traffic flow across different operating sectors (with unoccupied capacities), even though they may require some more extra costs.

There also exists much less delay assigned with the alternatives than the initials, which is resulted from the quota setting (for incentive concern) as discussed in Sec. 5.4. Then, among the different delay preferences, ground holding is obviously the most-commonly used, absorbing almost all the required system delays, but it appears that a small amount of airborne holding (27 min) and linear holding (46 min) could contribute to minimizing the total cost even if their unit cost of 90 Euro/min is higher than the cost of ground holding (81 Euro/min).

In terms of delay recovery, besides reducing the system cost explicitly, it also has some similar effects to air holding. For example, some available capacity (or free slot), which results from delaying a specific flight, might be taken over by another flight that is capable of advancing its arrival time at that place (through performing delay recovery), which is similar to an intermediate slot swapping process.

5.7. Trade-off of delay cost with extra fuel consumption and route charges

Benefiting from the accurate avoidance information for individual flights (see Sec. 3.3), the alternative trajectory that is precisely re-designed by AUs may incur as little extra costs as possible (compared to the initially scheduled). The distributions of extra fuel consumption and extra route charges for all the submitted lateral (i.e., 1,305) and vertical (i.e., 1,379) alternative trajectories are presented in Figs. 8c and 8d. On the other side, Fig. 11 shows the case for the selected trajectories, in which the trade-off of such extra trajectory cost with respect to the delay cost reduction (with respect to the GH mode) is given for each flight.

Among the set of submitted trajectories, there are some cases where a large amount of extra fuel consumption and/or route charges are needed (due to certain flights having multiple sectors to avoid for instance). Yet, it does not mean that these costly trajectories will be selected by the NM (based on the DCB model), as shown in Fig. 11 (in which only three flights are observed to incur an extra cost of higher than 1,000 Euro). Most costs, however, lie within an area of less than 300 Euro. Compared with this small amount of extra trajectory cost, the benefits gained from delay cost reduction is significant. Such advantage would be an explicit incentive that motivates AUs to use the proposed framework. Nevertheless, it can be noticed that a few flights somehow have negative (little though) net benefits (see Fig. 11), which also raises the necessity of taking into account more fairness concerns in future work.

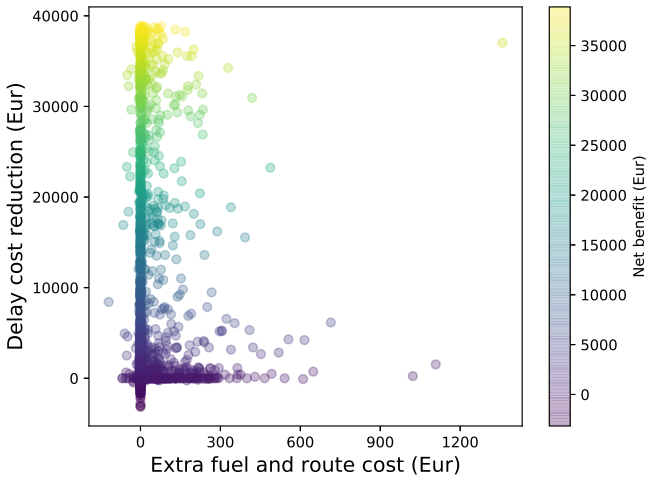


Figure 11: Trade-off between delay cost reduction with respect to C-ATFM (GH mode) and cost incurred from extra fuel consumption and route charges for selecting alternative trajectories.

6. Conclusions

This paper presented an innovative collaborative air traffic flow management framework in the scope of the future trajectory based operations. The main contribution of this paper is the integration of trajectory planning and air traffic flow management, as well as their interactive processes (i.e., collaborative trajectory design including hotspot detection and avoidance), into a single framework, and also the application to a real-world scale problem. The key finding of the paper is to quantify the benefit pool that might be attained with such framework operating in a deterministic setting.

The interactive trajectory design process (between airspace users and the Network Manager) was regarded as the key enabler of the framework for a series of downstream performance enhancements. Specifically, the accurate provision of the identified time-varying hotspot airspaces contributed to assisting airspace users to schedule alternative trajectories with as few extra costs as possible. Combining different options, resulted from the collaborative process, to manage the imbalances of demand and capacity could improve the cost-efficiency of air traffic flow management.

The linear optimization model, taking fully advantage of the trajectory design solutions, incorporated potential traffic management initiatives (i.e.,

multiple trajectory options mixed with different pre-tactical delay management preferences) to balance the traffic demand with airspace capacity, trying to minimize the deviation to the initial set of user-preferred trajectories. Results showed that, by implementing the proposed framework, a significant delay reduction was achieved with only a relatively small amount of extra costs (including fuel consumption and route charges) incurred.

Nevertheless, the achieved results represent an upper bound of the potential gains that the framework can provide, which will largely depend on the level of uncertainty in the system. A follow-up work will be conducted to further assess such impacts by adapting to the stochastic models proposed in the appendix. In addition to the analysis of the equity metrics, a more mature mechanism should be addressed to better balance the efficiency and equity. The computational performance will be also under future research towards an operational decision-support tool.

Acknowledgment

The work presented in this paper was partially funded by grants from the Funds of China Scholarship Council (No. 201506830050) and by the SESAR Joint Undertaking under grant agreement No. 699338, as part of the European Unions Horizon 2020 research and innovation programme: APACHE project (Assessment of Performance in current ATM operations and of new Concepts of operations for its Holistic Enhancement - http://apache-sesar.barcelonatech-upc.eu/en). The opinions expressed herein reflect the authors view only. Under no circumstances shall the SESAR Joint Undertaking be responsible for any use that may be made of the information contained herein. The authors would like to thank Dr. Christina Barado at UPC for generating the experimental results with the CASA algorithm. The authors appreciate the reviewers and the editor for all their valuable comments and suggestions.

Appendix A Notations

Acronym:

AD	Arrival Delay
AH	Airborne Holding
ATC	Air Traffic Control
ATFM	Air Traffic Flow Management
ATM	Air Traffic Management
ATS	Air Traffic Service
AUs	Airspace Users
CASA	Computer Assisted Slot Allocation
C-ATFM	Collaborative Air Traffic Flow Management
CDM	Collaborative Decision-Making
CTA	Controlled Time of Arrival
CTD	Controlled Time of Departure
CTO	Controlled Time Over
DCB	Demand and Capacity Balancing
DDR2	Demand Data Repository v2
DR	Delay Recovery
ETA	Estimated Time of Arrival
ETD	Estimated Time of Departure
ETO	Estimated Time Over
GCD	Great Circle Distance
GH	Ground Holding
LH	Linear Holding
NM	Network Manager
O/D	Origin and Destination
RBS	Ration-by-Schedule
SR	Structured Route
TBO	Trajectory Based Operations

Nomenclature:

$f \in \mathcal{F}$	set of flights
$k \in \mathcal{K}$	set of trajectories
$j \in \mathcal{J}$	set of control points
$t \in \mathcal{T}$	set of time moments
$l \in \mathcal{L}$	set of operating sectors

$\tau \in \mathbb{T}$	set of time periods
$\xi \in \Xi$	set of scenarios
$s \in \mathcal{S}$	set of stages
$\mathcal{K}_f$	subset of trajectory options of flight f
$\mathcal{T}_k^j$	subset of feasible time window for trajectory k at position j
$\mathcal{L}_\tau$	subset of operating sectors that are open in time period τ
$\mathcal{J}_a$	subset of airports
$\mathcal{J}_l^\tau$	subset of elementary sectors collapsing to operating sector l during time period τ
$\mathcal{J}_k$	subset of control points that trajectory k traverses
$\mathcal{J}_k(i)$	$\begin{cases} \text{departure airport,} & \text{if } i = 1 \\ \text{arrival airport,} & \text{if } i = n_k \\ \text{intermediate designed positions,} & \text{if } 1 < i < n_k \end{cases}$
$\mathcal{T}(\tau)$	subset of time moments subject to time period τ
$\mathcal{T}_s$	subset of time moments subject to stage s
$\mathcal{S}_\xi$	subset of stages subject to scenario ξ
$\mathcal{F}_s$	subset of flights initially scheduled to depart within stage s
N_t	number of trajectory options submitted for flight f
N_d	number of pre-tactical delay preferences submitted for flight f
N_f	number of possible alternatives for flgiht f to solve the DCB problem
r_k^j	initially scheduled time of trajectory k at position j
$\underline{\mathcal{T}}_k^j$	lower bound of the feasible time window $\mathcal{T}_k^j$
$\overline{\mathcal{T}}_k^j$	upper bound of the feasible time window $\mathcal{T}_k^j$
F_k	fuel consumption of trajectory k
F_0^f	initially scheduled fuel consumption for flight f
R_k	route charges of trajectory k
R_0^f	initially scheduled route charges for flight f
ξ_s^n	nth scenario ξ that may occur at the end of stage s
$z_k^{j,j'}$	scheduled segment flight time of two contiguous positions j and j'
$u_k^{j,j'}$	time bound of delay recovery within flight segment (j, j')
$v_k^{j,j'}$	time bound of linear holding within flight segment (j, j')
$S(k, l, \tau)$	the first entered elementary sector for trajectory k among those collapsed into operating sector l during time period τ
$C_j^D(\tau)$	airport departure capacity during time period τ
$C_j^A(\tau)$	airport arrival capacity during time period τ
$C_l(\tau)$	capacity of operating sector l during time period τ
α_k	weighting cost of extra fuel consumption for trajectory k
β_k	weighting cost of extra route charges for trajectory k

γ_k^i	weighting cost of types of delays for trajectory k, i.e., ground holding $i = g$, air holding $i = a$ and delay recovery $i = d$
ϵ	fairness factor for equivalent delay assignment
$P(\xi)$	probability of scenario ξ that occurs
$\dot{T}_k^j$	assigned time for trajectory k departing from position j

Appendix B Establishing sector scheme from published database

The following Algorithm 2 presents a procedure to establish the static sector scheme according to the DDR2 database, where the required source files include: 1) OpeningScheme.cos, 2) Configuration.cfg, and 3) Airspace.spc, for the same AIRAC date. As for mapping the operating capacities to the set of matched collapsed sectors, some extra sources are needed, such as 4) TrafficVolume.ntfv, 5) Activation.nact, and 6) Capacity.ncap (see (Xu and Prats, 2017b) for a detailed illustration to the procedure).

Algorithm 2 Retrieve static collapsing scheme for elementary sectors

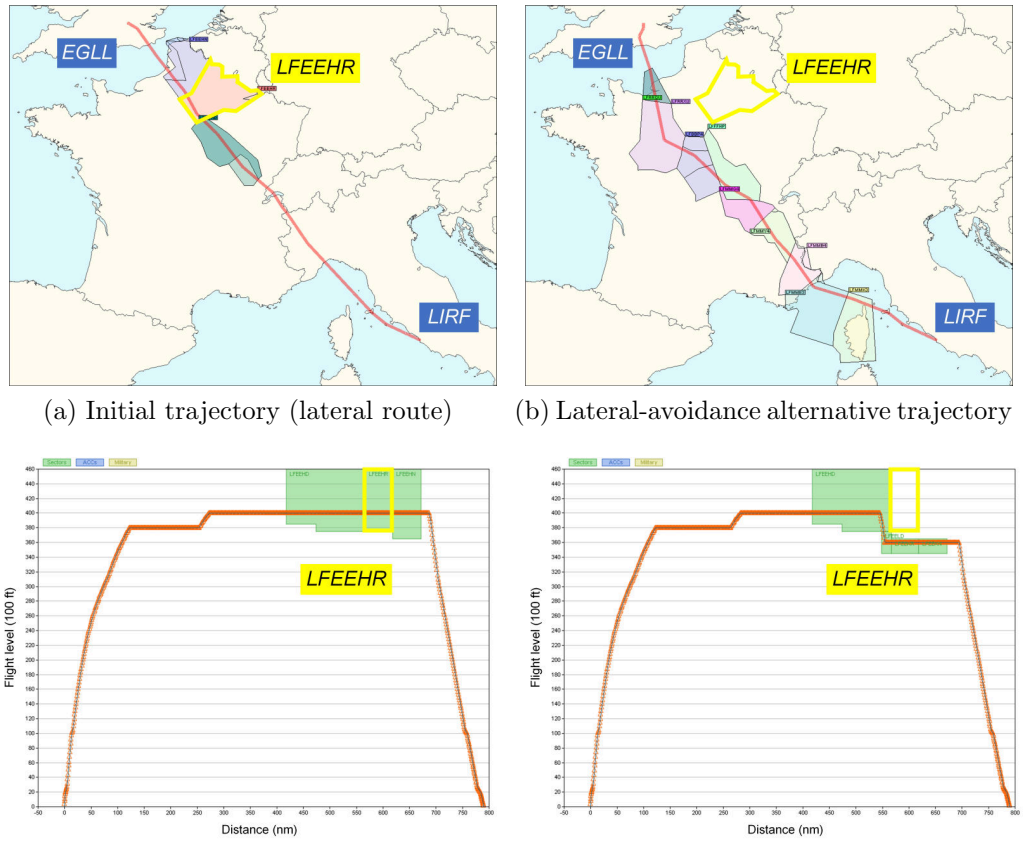
```
for e in elementary_sector_list do
    for t in time_period_list do
        for a in area_control_center_opening_list do
            for cf in configuration_list[a][t] do
                for s in operating_sector_list[cf] do
                    if s in elementary_sector_list then
                        if s == e then
                            collapse_scheme[e][t] = s
                    else if s in collapsed_sector_list then
                        for c in collapsed_sector_list[s] do
                            if c == e then
                                collapse_scheme[e][t] = s
```

Appendix C Extended models concerning uncertainty

See online supplement materials.

Appendix D Involvement of a specific flight in the framework

This appendix presents the involvement of airspace users in the interactive procedures in collaboration with the Network Manager, under the proposed C-ATFM framework. It is shown via an example of a specific flight (LIRF-EGLL) extracted from the scenario simulated for the numerical experiments in Sec. 5.



(a) Initial trajectory (lateral route) (b) Lateral-avoidance alternative trajectory

(c) Initial trajectory (vertical profile) (d) Vertical-avoidance alternative trajectory

Figure 12: Airspace user's submitted trajectory options for the specific flight (LIRF-EGLL), including the initially scheduled trajectory, the lateral- and vertical-avoidance alternative trajectories.

Fig. 12a shows the lateral route of the initial trajectory. As discussed in Sec. 3.1, the initial trajectory should reflect the most-preferred trajectory, minimizing the aircraft direct operating cost, for the airspace user. The

traversed elementary sectors are highlighted in the two figures, whilst the vertical profile of the trajectory intersecting with those sectors is as shown in Figs. 12c. For each of the traversed elementary sectors, the control point is defined at the entry position. Through a primary hotspot detection process (recall Sec. 3.2) conducted by the Network Manager, comparing the traffic demand (of the initial trajectories) and airspace capacity for different time periods, sector LFEEHR (see Figs. 12a and 12c) is identified as the hotspot area to be avoided for this particular flight.

Table 6: Costs of all trajectory options for the specific flight (LIRF-EGLL).

Trajectory	**Fuel (kg)**	**Charges (Euro)**	**Ex.fuel (kg)**	**Ex.charges (Euro)**
Initial	4,529	1,359	0	0
Lateral	4,666	1,370	137	11
Vertical	4,548	1,359	19	0

With the detailed avoidance information received (recall Sec. 3.3), the airspace user then produces the lateral hotspot-avoidance (see Fig. 12b) and vertical hotspot-avoidance (see Figs. 12d) alternative trajectories to precisely evade the corresponding sectors, using the aircraft trajectory optimization techniques introduced in Sec. 3.4.

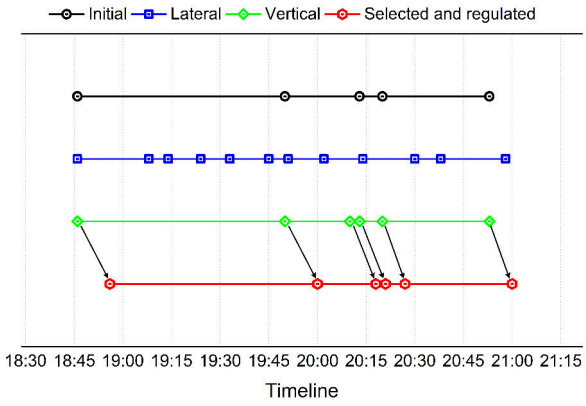


Figure 13: Timeline of each trajectory along different defined positions (i.e., origin and destination airport, as well as entry point of all elementary sectors the trajectory traverses).

Taking the extra cost into account (see Table 6), the airspace user decides to submit all the three trajectory options to the Network Manager, as well as the preferences for different types of delay (recall Sec. 3.5).

Subsequently, the DCB optimization model is initiated by the Network Manager, generating the optimal solution for trajectory selections and delay assignments, in such a way to minimize the overall deviation with respect to the airspace users' initial trajectories (recall Sec. 4). Eventually, the vertical hotspot-avoidance alternative trajectory is selected for the particular flight (see green line in Fig. 13), based on which 10 min of ground holding (i.e., 150 Euro of extra delay cost) is imposed and 3 min of delay recovery (i.e., -15 Euro of extra delay cost) is allowed, leading consequently to 7 minutes of arrival delay.

Appendix E Sensitivity analysis for cost parameters

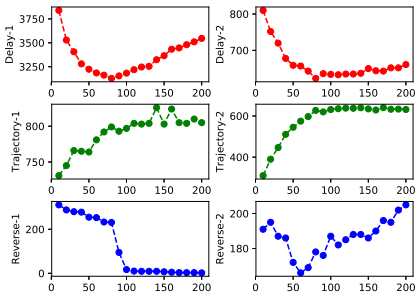
Additional experiments have been conducted to assess the benefit pool of using the proposed framework. The impacts of some independent variables are considered including the cost of delay (by means of ground holding, air holding and delay recovery), the non-linearity of delay cost (in particular the final arrival delay), and the quota of delay for alternative trajectories (as introduced in Sec. 5.4). Table 7 presents these variables and their associated ranges taken in the sensitivity study.

Table 7: Independent variables in the sensitivity analysis.

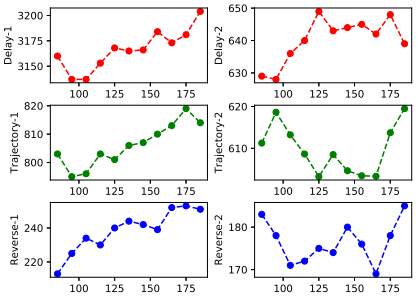
Variables	Baseline	Min	Max	Step	Num
Ground holding unit cost (Euro/min)	81	10	200	10	20
Air holding unit cost (Euro/min)	90	85	185	10	11
Delay recovery unit cost (Euro/min)	-5	-80	0	10	9
Non-linearity delay cost factor (ϵ)	0.2	0	1	0.1	11
Quota of delay for alternatives (min)	10	0	45	5	10

In terms of the effects of ground holding unit cost, the results are as shown in Fig. 14a. It is obvious that large amount of delay is issued when the unit cost of ground holding is relatively low, and many flights are involved to share this delay, which yields a high number of departure reversal pairs. When the unit cost of ground holding is greater than that of air holding (i.e., 90 Euro/min), a sharp decrease occurs in the departure reversals as ground holding is less preferred (which increases the total delay though). Yet, the number of arrival reversals keeps increasing due to the fact that more air holding is performed.

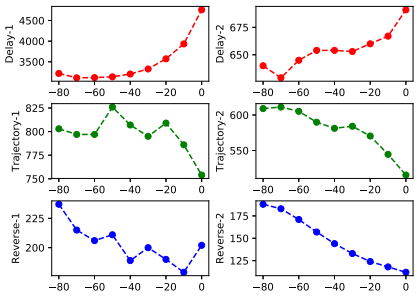
For the unit cost of air holding (see Fig. 14b), as it becomes more expensive the total system delay increases (through distributing more ground



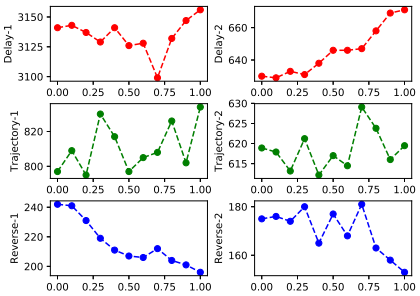
(a) Ground holding unit cost



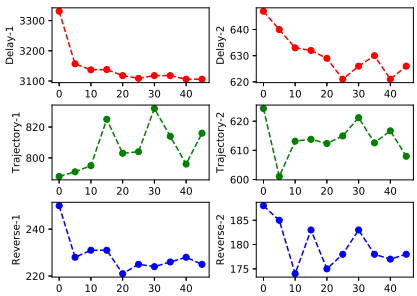
(b) Air holding unit cost



(c) Delay recovery unit cost



(d) Non-linearity delay cost factor



(e) Quota of delay for alternative

Figure 14: Sensitivity results. Delay-1: Total delay (min); Delay-2: Total delayed flights (#); Trajectory-1: Selected alternative trajectories (#); Trajectory-2: Extra trajectory cost (100 Euro); Reverse-1: Departure reversals (#); Reverse-2: Arrival reversals (#)

holding), which shows the benefits of including air holding in the delay assignment. In the meantime, the number of flights selecting alternative trajectories has formed a positive relation with the cost of air holding, whilst the extra trajectory cost turns to be not relevant.

Then, with regards to the delay recovery unit cost, as shown in Fig. 14c, it turns out that the more benefits the delay recovery could raise (e.g., -80 Euro/min means that recovering 1 min of delay could generate 80 Euro benefits), the less system delay is eventually realised (as most of the delay caused by ground and air holding can be recovered). In accordance with the higher amount of delay recovered, the flight reversal pairs on arrival increase as well, due to the more often speed (and thus sequence) changes en route.

The factor of non-linearity in delay cost also has certain effects to the results (see Fig. 14d). Generally, the higher this non-linearity value is the more flights are involved in delay assignment (not necessarily for the total delay), meaning that the average delay for each flight can be lowered. In terms of fairness, as one would expect, the numbers of flight reversal pairs on both departure and arrival are reduced, due to the fact that delay is assigned moderately across more flights rather than unevenly to specific individuals.

Finally, Fig. 14e demonstrates the impacts of imposing a maximal delay bound to each alternative trajectory. This is to take AUs' incentives into account when submitting alternative trajectories for their flights. From a system point of view, this change does not necessarily worsen the global benefits, as long as such quota is not close to 0 min (which incurs higher delay, trajectory cost and reversal pairs at the same time).

References

Airlines for America, 2016. U.S. Passenger Carrier Delay Costs. Tech. rep. URL http://airlines.org/data/

Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B., 2010. Total delay impact study: a comprehensive assessment of the costs and impacts of flight delay in the United States. University of California, Berkeley. Institute of Transportation Studies.

Ball, M. O., Hoffman, R., Lovell, D., Mukherjee, A., 2005. Response mechanisms for dynamic air traffic flow management. In: Proceedings of the 6th USA/Europe ATM R&D Seminar. Baltimore. US.

Ball, M. O., Hoffman, R., Odoni, A. R., Rifkin, R., 2003. A stochastic integer program with dual network structure and its application to the ground-holding problem. Operations research 51 (1), 167–171.

Ball, M. O., Hoffman, R. L., Knorr, D., 2000. Assessing the benefits of collaborative decision making in air traffic management. In: Proceedings of the 3rd USA/Europe ATM R&D Seminar. Napoli, Italy.

Barnhart, C., Bertsimas, D., Caramanis, C., Fearing, D., 2012. Equitable and efficient coordination in traffic flow management. Transportation science 46 (2), 262–280.

Bertsimas, D., Farias, V. F., Trichakis, N., 2011a. The price of fairness. Operations research 59 (1), 17–31.

Bertsimas, D., Gupta, S., 2015. Fairness and collaboration in network air traffic flow management: an optimization approach. Transportation Science 50 (1), 57–76.

Bertsimas, D., Lulli, G., Odoni, A., 2011b. An integer optimization approach to large-scale air traffic flow management. Operations research 59 (1), 211–227.

Bertsimas, D., Patterson, S. S., 1998. The air traffic flow management problem with enroute capacities. Operations research 46 (3), 406–422.

Bertsimas, D., Patterson, S. S., 2000. The traffic flow management rerouting problem in air traffic control: A dynamic network flow approach. Transportation Science 34 (3), 239–255.

Betts, J., 2010. Practical Methods for Optimal Control and Estimation Using Nonlinear Programming, 2nd Edition. Society for Industrial and Applied Mathematics.
URL https://epubs.siam.org/doi/abs/10.1137/1.9780898718577

Betts, J. T., Cramer, E. J., 1995. Application of direct transcription to commercial aircraft trajectory optimization. Journal of Guidance, Control, and Dynamics 18 (1), 151–159.

Bolić, T., Castelli, L., Corolli, L., Rigonat, D., 2017. Reducing atfm delays through strategic flight planning. Transportation Research Part E: Logistics and Transportation Review 98, 42–59.

Chang, K., Howard, K., Oiesen, R., Shisler, L., Tanino, M., Wambsganss, M. C., 2001. Enhancements to the FAA ground-delay program under collaborative decision making. Interfaces 31 (1), 57–76.

Cook, A., 2007. European air traffic management: principles, practice, and research. Ashgate Publishing, Ltd.

Cook, A. J., Tanner, G., 2015. European airline delay cost reference values: Updated and extended values. Tech. rep.

Dalmau, R., Melgosa, M., Vilardaga, S., Prats, X., 2018. A fast and flexible aircraft trajectory predictor and optimiser for atm research applications. In: Proceedings of the 8th International Conference for Research in Air Transportation (ICRAT). Castelldefels, Spain.

Dalmau, R., Prats, X., 2017. Assessing the impact of relaxing cruise operations with a reduction of the minimum rate of climb and/or step climb heights. Aerospace Science and Technology 70, 461–470.

Delgado, L., 2015. European route choice determinants. In: Proceedings of the 11th USA/Europe ATM R&D Seminar. Lisbon, Portugal.

EUROCONTROL, 2017a. All-causes delay and cancellations to air transport in Europe. Tech. Rep. CODA Digest 2016, CDA-2017-005, Network Manager.

EUROCONTROL, 2017b. ATFCM operations manual - network operations handbook. Tech. Rep. Ed. 21.0.

EUROCONTROL, 2018. DDR2 reference manual for general users. Tech. Rep. Version 2.9.5.

FAA, 2009. Traffic Flow Management in the National Airspace System. Tech. Rep. FAA-2009-AJN-251, Federal Aviation Administration.

FAA, 2014. Collaborative Trajectory Options Program (CTOP): Document Information. Tech. Rep. AC 90-115, Federal Aviation Administration.

FAA, 2018. Nextgen implementation plan. Tech. Rep. 2018-2019 Edition.

Gardi, A., Sabatini, R., Ramasamy, S., 2016. Multi-objective optimisation of aircraft flight trajectories in the ATM and avionics context. Progress in Aerospace Sciences 83, 1–36.

Hart, P. E., Nilsson, N. J., Raphael, B., 1968. A formal basis for the heuristic determination of minimum cost paths. IEEE transactions on Systems Science and Cybernetics 4 (2), 100–107.

Holloway, S., 2008. Straight and level: Practical airline economics. Ashgate Publishing, Ltd.

ICAO, 2016. Global air navigation plan (2016-2030), Doc 9750-an/963, fifth edition. Tech. rep.

Kim, A., Hansen, M., 2015. Some insights into a sequential resource allocation mechanism for en route air traffic management. Transportation Research Part B: Methodological 79, 1–15.

Libby, M., Buckner, J., Brennan, M., 2005. Operational concept for Airspace Flow Programs (AFP). Tech. rep., FAA Air Traffic Organization, Systems Operations Services.

Liu, Y., Hansen, M., Lovell, D., Ball, M., 2018. Predicting aircraft trajectory choice - a nominal route approach. In: Proceedings of the 8th International Conference for Research in Air Transportation (ICRAT). Castelldefels, Spain.

Lulli, G., Odoni, A., 2007. The European air traffic flow management problem. Transportation Science 41 (4), 431–443.

Meyer, J. R., Oster, C. V., 1981. Airline deregulation: the early experience. Auburn House.

Miller, M. E., Hall, W. D., 2015. Collaborative trajectory option program demonstration. In: Proceedings of the 34th IEEE/AIAA Digital Avionics Systems Conference (DASC). IEEE, Prague, Czech Republic, pp. 1C1–8.

Molina, M., Carrasco, S., Martin, J., 2014. Agent-based modeling and simulation for the design of the future european air traffic management system: the experience of cassiopeia. In: International Conference on Practical Applications of Agents and Multi-Agent Systems. Springer, pp. 22–33.

Morrison, S., Winston, C., 2010. The economic effects of airline deregulation. Brookings Institution Press.

Mukherjee, A., Hansen, M., 2009. A dynamic rerouting model for air traffic flow management. Transportation Research Part B: Methodological 43 (1), 159–171.

Nuic, A., Mouillet, V., 2014. User Manual for the Base of Aircraft Data (BADA) Family 4, EEC Technical/Scientific Report. Eurocontrol Experimental Centre, Bretigny-sur-Orge, France.

Odoni, A. R., 1987. The flow management problem in air traffic control. In: Flow control of congested networks. Springer, Berlin, Heidelberg, pp. 269–288.

Pourtaklo, N. V., Ball, M., 2009. Equitable allocation of enroute airspace resources. In: Proceedings of the 8th USA/Europe ATM R&D Seminar. Napa, CA, US.

Richetta, O., Odoni, A. R., 1994. Dynamic solution to the ground-holding problem in air traffic control. Transportation research part A: Policy and practice 28 (3), 167–185.

Rios, J., Ross, K., 2010. Massively parallel Dantzig-Wolfe decomposition applied to traffic flow scheduling. Journal of Aerospace Computing, Information, and Communication 7 (1), 32–45.

Rodionova, O., Arneson, H., Sridhar, B., Evans, A., Sept 2017. Efficient trajectory options allocation for the ollaborative trajectory options program. In: Proceedings of the 36th IEEE/AIAA Digital Avionics Systems Conference (DASC). St. Petersburg, FL, US, pp. 1–10.

SESAR, 2015. Step 1 v3 UDPP validation report, optimised airspace user operations. Tech. Rep. PJ07.06.02, SESAR JU.

SESAR, 2017. SESAR 2020 concept of operations edition 2017. Tech. Rep. PJ.19-02.

SESAR, 2020. European ATM master plan. Tech. Rep. 2020 Edition.

Sherali, H. D., Hill, J. M., McCrea, M. V., Trani, A. A., 2011. Integrating slot exchange, safety, capacity, and equity mechanisms within an airspace flow program. Transportation science 45 (2), 271–284.

Soler, M., Olivares, A., Staffetti, E., Zapata, D., 2012. Framework for aircraft trajectory planning toward an efficient air traffic management. Journal of Aircraft 49 (1), 341–348.

Tandale, M. D., Wiraatmadja, S., Vaddi, V. V., Rios, J. L., 2013. Massively parallel optimal solution to the nationwide traffic flow management problem. In: 2013 Aviation Technology, Integration, and Operations Conference. Los Angeles, CA, US, p. 4349.

Terrab, M., Paulose, S., 1992. Dynamic strategic and tactical air traffic flow control. In: IEEE International Conference on Systems, Man and Cybernetics. IEEE, Chicago, US, pp. 243–248.

US Department of Transportation, 2016. Airline on-time statistics. Tech. rep. URL https://www.transtats.bts.gov/

Vossen, T., Ball, M., 2006a. Optimization and mediated bartering models for ground delay programs. Naval Research Logistics (NRL) 53 (1), 75–90.

Vossen, T. W., Ball, M. O., 2006b. Slot trading opportunities in collaborative ground delay programs. Transportation Science 40 (1), 29–43.

Vossen, T. W. M., Hoffman, R., Mukherjee, A., 2012. Air Traffic Flow Management. Springer US, Boston, MA, pp. 385–453.

World Meteorological Organization, 1994. A guide to the code form fm 92-ix ext. GRIB. Edition 1.

Xiong, J., Hansen, M., 2009. Value of flight cancellation and cancellation decision modeling: ground delay program postoperation study. Transportation Research Record: Journal of the Transportation Research Board (2106), 83–89.

Xu, Y., Dalmau, R., Prats, X., 2017. Maximizing airborne delay at no extra fuel cost by means of linear holding. Transportation Research Part C: Emerging Technologies 81, 137–152.

Xu, Y., Prats, X., 2017a. Effects of linear holding for reducing additional flight delays without extra fuel consumption. Transportation Research Part D: Transport and Environment 53, 388–397.

Xu, Y., Prats, X., 2017b. Including linear holding in air traffic flow management for flexible delay handling. Journal of Air Transportation 25, 123–137.

Xu, Y., Prats, X., Delahaye, D., 2018. Synchronization of traffic flow and sector opening for collaborative demand and capacity balancing. In: 37th IEEE/AIAA Digital Avionics Systems Conference (DASC). IEEE, London, UK.

Yang, Y., 2017. Practical method for 4-dimentional strategic air traffic management problem with convective weather uncertainty. IEEE Transactions on Intelligent Transportation Systems 19 (6), 1697–1708.

Zelinski, S., Lai, C. F., 2011. Comparing methods for dynamic airspace configuration. In: Proceedings of the 30th IEEE/AIAA Digital Avionics Systems Conference (DASC). IEEE, Seattle, WA, US, pp. 3A1–1.

2020-03-03

A framework for collaborative air traffic flow management minimizing costs for airspace users: enabling trajectory options and flexible pre-tactical delay management

Xu, Yan

Elsevier

Xu Y, Dalmau R, Melgosa M, et al., (2020) A framework for collaborative air traffic flow management minimizing costs for airspace users: enabling trajectory options and flexible pre-tactical delay management. Transportation Research Part B: Methodological, Volume 124, April 2020, pp. 229-255

https://doi.org/10.1016/j.trb.2020.02.012